SEP
SECRETARÍA DE EDUCACIÓN PÚBLICA

Instituto Politécnico Nacional
"La Técnica al Servicio de la Patria"

ISSN: 1870-4069

RCS

Research in Computing Science

Vol. 81

Advances in Computing Science

L. E. Berthet Dorantes
J. C. Chimal Eguía
E. A. Santos Camacho
E. Castillo Montiel

Advances in Computing Science

Research in Computing Science

Series Editorial Board

Editors-in-Chief:

Grigori Sidorov (Mexico)
Gerhard Ritter (USA)
Jean Serra (France)
Ulises Cortés (Spain)

Associate Editors:

Jesús Angulo (France)
Jihad El-Sana (Israel)
Jesús Figueroa (Mexico)
Alexander Gelbukh (Russia)
Ioannis Kakadiaris (USA)
Serguei Levachkine (Russia)
Petros Maragos (Greece)
Julian Padget (UK)
Mateo Valero (Spain)

Editorial Coordination:

María Fernanda Rios Zacarias

Formatting:

Luis Enrique Berthet Dorantes
Juan Carlos Chimal Eguía
Evelia Araceli Santos Camacho
Ana Bertha Cruz Martínez
Erandi Castillo Montiel
Iliac Huerta Trujillo
Jesús Emmanuel Velázquez Cruz
Felipe de Jesús Lara Vargas
Mauricio Iván Guerrero Hernández
Luis Alberto Ortiz Chan
Armando Castañeda Tenango
Eder Sánchez Díaz

Research in Computing Science es una publicación trimestral, de circulación internacional, editada por el Centro de Investigación en Computación del IPN, para dar a conocer los avances de investigación científica y desarrollo tecnológico de la comunidad científica internacional. **Volumen 81,** noviembre 2014. Tiraje: 500 ejemplares. *Certificado de Reserva de Derechos al Uso Exclusivo del Título* No. : 04-2005-121611550100-102, expedido por el Instituto Nacional de Derecho de Autor. *Certificado de Licitud de Título* No. 12897, *Certificado de licitud de Contenido* No. 10470, expedidos por la Comisión Calificadora de Publicaciones y Revistas Ilustradas. El contenido de los artículos es responsabilidad exclusiva de sus respectivos autores. Queda prohibida la reproducción total o parcial, por cualquier medio, sin el permiso expreso del editor, excepto para uso personal o de estudio haciendo cita explícita en la primera página de cada documento. Impreso en la Ciudad de México, en los Talleres Gráficos del IPN – Dirección de Publicaciones, Tres Guerras 27, Centro Histórico, México, D.F. Distribuida por el Centro de Investigación en Computación, Av. Juan de Dios Bátiz S/N, Esq. Av. Miguel Othón de Mendizábal, Col. Nueva Industrial Vallejo, C.P. 07738, México, D.F. Tel. 57 29 60 00, ext. 56571.

Editor responsable: *Grigori Sidorov*, *RFC SIGR651028L69*

Research in Computing Science is published by the Center for Computing Research of IPN. **Volume 81**, Noviembre 2014. Printing 500.

The authors are responsible for the contents of their articles. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior permission of Centre for Computing Research. Printed in Mexico City, in the IPN Graphic Workshop – Publication Office.

Volume 81

Advances in Computing Science

Luis Enrique Berthet Dorantes,
Juan Carlos Chimal Eguía,
Evelia Araceli Santos Camacho,
Erandi Castillo Montier (eds.)

Instituto Politécnico Nacional, Centro de Investigación en Computación
México 2014

ISSN: 1870-4069

Copyright © Instituto Politécnico Nacional 2014

Instituto Politécnico Nacional (IPN)
Centro de Investigación en Computación (CIC)
Av. Juan de Dios Bátiz s/n esq. M. Othón de Mendizábal
Unidad Profesional "Adolfo López Mateos", Zacatenco
07738, México D.F., México

http://www.rcs.cic.ipn.mx
http://www.ipn.mx
http://www.cic.ipn.mx

The editors and the publisher of this journal have made their best effort in preparing this special issue, but make no warranty of any kind, expressed or implied, with regard to the information contained in this volume.

All rights reserved. No part of this publication may be reproduced, stored on a retrieval system or transmitted, in any form or by any means, including electronic, mechanical, photocopying, recording, or otherwise, without prior permission of the Instituto Politécnico Nacional, except for personal or classroom use provided that copies bear the full citation notice provided on the first page of each paper.

Indexed in LATINDEX and Periodica / Indexada en LATINDEX y Periódica

Printing: 500 / Tiraje: 500

Printed in Mexico / Impreso en México

Preface

The purpose of this volume is to present the most recent advance in selected areas of Computer Science, The works included in this volume where carefully selected by the editors on the basis of the blind reviewing process and the main criteria for selection were originality and technical quality. This issue of Research in Computing Science will be useful for researches and students working in he different areas of Computer Science, as well as, for all reader interested you want to enrich your knowledge in this file.

In total, we received 72 paper that were submitted for evaluation; each submitted paper was reviewed by 2 independent members of the editorial board of the volume or additional reviewers. The acceptance rate is 38%. This volume contains revised version of 16 accepted papers, divided into four sections corresponding to the areas: Artificial Intelligence, Communications & Computer Networks, Databases & Software Technology and Digital Signal Processing. The missing papers are published in the volume number 83 of Research in Computing Science.

We would like express our gratitude to all people who help to elaborate this volume. First to the authors of the papers for the technical excellence of their works, that guarantees the quality of this publications. We also want to thank the members of the editorial board for their hard work in evaluating and selecting the best's papers out of many submissions that we received. We express sincerely our gratitude to the Sociedad Mexicana de Inteligencia Artificial (SMIA) for its collaboration in elaborates this publication. Also we want to give special recognition to the Centro de Investigación en Computacion of the Instituto Politécnico Nacional (CIC-IPN) for facilities give in order to achieve the success in the publication of this volume. The submission, reviewing, and selection process was supported for free by the EasyChair system, www.EasyChair.org. Also we want to give special recognition to ORACLE, its participation as a sponsor.

Luis Enrique Berthet Dorantes
Juan Carlos Chimal Eguía
Evelia Araceli Santos Camacho
Erandi Castillo Montiel
November 2014

Table of Contents

Digital Signal Processing

Artificial Intelligence

Rehabilitation games with active markers for children with infantile cerebral palsy

Lee Wan Chen[1], Rubén Posada Gómez[2,] Ulises Juárez-Gómez[3], Guillermo C. Robles[4]

Intituto Tecnologico de Orizaba, Mexico

[1]leewc5689@gmail.com,[2]pgruben@yahoo.com,[3]ujuarez@ito-depi.edu.mx,[4]gc_robles@hotmail.com

Abstract. This paper shows a combination of image processing technology and active markers in a medical application. Developing a human-computer interface prototype, the proposed system uses a PlayStation 3™ EyeCamera and active marker to perform neuromuscular rehabilitation games for upper limbs, as a tool to support rehabilitation therapies. As a case of study, we observed children with infantile cerebral palsy playing the rehabilitation games prototype during six months, obtaining a quite positive result. This paper shows a new low cost, attractive and effective tool, allowing multiple therapies of neuromuscular rehabilitation to satisfy different needs of patients.

Keywords: Rehabilitation, cerebral palsy, computer games, image processing, active marker.

1 Introduction

Today, the interaction between humans and computers becomes more and more common to support people's activities in a more natural way, giving more uses in different areas. Such as education [1], work [2], home [3], medicine [4], among others. The technology of human - computer integration or integration of robot and computer are used in rehabilitation of the medicine area [5,6] in the way it can achieve distinct type of rehabilitation. It is more attractive to patients performing tedious and repetitive exercises [7], on the other hand; it is also more economic compared to the traditional rehabilitation devices and allows patients to do their therapy at home. For all these advantages, computer is an excellent tool to support rehabilitation. [8] Shows that virtual reality technology offers an efficient way to help rehabilitation.

Currently there is a consensus, considering the cerebral palsy (CP) as a group of developmental disorders of movements and posture, causing activity limitations, which are attributed to a non-progressive attack on the developing brain in the fetal period and early years. From 1000 births, approximately 2 or 3 have CP [9].

According to the problem of children with cerebral palsy, a good technique of rehabilitation therapies will be very important. This paper shows the development of a human computer interface using image processing with a PlayStation 3TM EyeCame-

©L. Berthet, J. C. Chimal, E. A. Santos, E. Castillo
Advances in Computer Science
Research in computing Science 81, 2014 pp. 3-12

ra and a LED pen as an active marker. The system includes the development of neuromuscular rehabilitation games for upper limbs, in order to help children with infantile cerebral palsy on his therapy. And finally, this paper shows the result of case studies with children with infantile cerebral palsy.

2 Neuromuscular disease

Neuromuscular diseases are characterized by muscle weakness, which could result in difficulty of limbs movements and a decrease of the functional ability [10]. But it also, depending on the distribution of the weakness, difficulty in swallowing, chewing, speaking or breathing. Also secondary complications may occur; for example: contractures and scoliosis, which may just them, incapacitate the patient.

A relevant muscular and functional evaluation is necessary to document the natural history of the patient and then determine the outcome of the therapy.

The main treatment of neuromuscular diseases can be summarized in:

1. Maintenance or improvement of muscle strength.
2. Prevention of contractures or deformities.
3. Maintenance or improvement of function.
4. Stimulation or prolongation of ambulation. [10]

3 Human-Computer Interface with games of rehabilitation neuromuscular

3.1 Installing the device

To achieve the natural manipulation of the human-computer interface that we present requires a large computer screen or a TV or a projector connected to the computer, a webcam that detects all area of the screen and an active marker to manipulate the screen. In this case we use PlayStation 3TM EyeCamera as a webcam, because it presents a faster response time of the image than a normal web camera and the detection of image is more stable. We also use a projector and a color LED pen (red LED, green LED or blue LED) as active marker to manipulate the interface, such as shown in Figure 1.

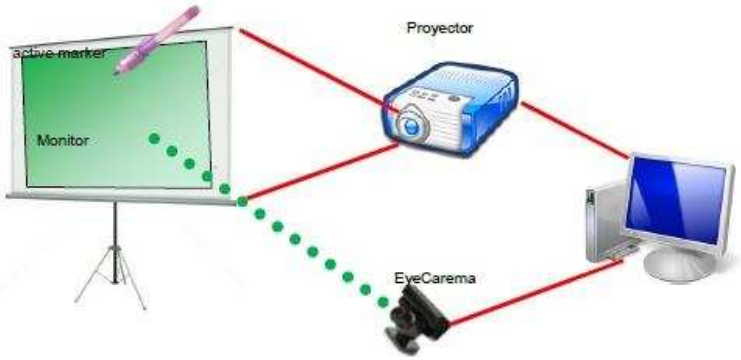


Fig. 1. Installing the device

3.2 Marker

The marker can be a device of control or an object that manipulates the software with another device capturing signals of marker and transmitting the signal to the computer. Currently there are different types of control devices; generally we can divide them into 2 types: active markers (which send signals to the device) and passive markers (which do not emit the signal, and wait until the device finds them).

In this project we developed some active markers, and we compared their advantage and disadvantage. We design glove with infrared LED, hook and loop fastener ring with an infrared LED, pen with color LED and hook and loop fastener bracelet with color LED. Here we use hook and loop fastener to adjust the size of the wrist and finger of the user. See Figure 2

Fig. 2. Active markers

When comparing the detection result of infrared LED marker and the LED color marker, it showed that to detect an infrared marker, a filter is needed to remove the

ambient lights, but it is not very effective. When the environment is a little lighter there may be too much intervention, if the infrared light is far, it wouldn't be detected. On the other hand, for the LED color marker, though we use a filter to remove most of the lights and colors, as this is search by specific LED color you can find less intervention, it doesn't depend much on ambient light like infrared. Also, in the future we'll be able to involve multi user games at the same time by having different LED colors.

Infrared LED and color LED are unidirectional, light is focused on the top, if the top of LED is turned in another direction, the light that is seen on the camera will be decreased a lot, it even disappears. This problem causes detection fault. To solve this problem we painted black color on the top, put LED inside a small sphere of glass and filled with hot melt, so that LED´s light will be in all different directions.

3.3 Image processing to detect active marker

First, it captures the image from the camera, then it gets the data from each pixel of the image in RGB format, and compares the data of each pixel with data in RGB of color LED light. When the color LED light is found in image, it returns his position x, y. Environment interference is strong enough to make the program get confused and it detects incorrect point. To avoid this problem we put a filter on the lens of EyeCamera. See Figure 3.

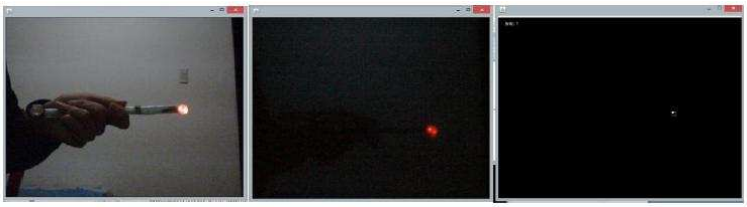


Fig. 3. Regular image with point detection filter.

3.4 Computer games for rehabilitation of upper limbs

We developed two games of neuromuscular rehabilitation for upper limb; car game rehabilitates continuo movement of upper limbs, and frog game rehabilitates movement of upper limb and movement of thumb.

Car game

This game works in following way: The car has to move in his road with the active marker. If marker is outside the road it will emit a sound that represents an error and errors committed during the game will be recorded. If the user puts the marker outside car's location it will emit another sound, which means it's not a valid movement. When user reaches the final goal sound of applause, it will mean that the user has

finished this level, the game saves automatically: game level, user name, date, and time that user takes to finish the game's level up to milliseconds and number of times that user spent outside of the road. The present game has various roads recommended by the doctor and divide levels by difficulty in order to help users to perform their rehabilitation and exercising different arm movements. See Figure 4.

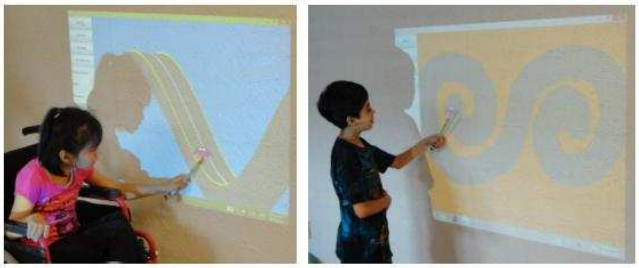


Fig. 4. Car game

In this program there is also a tool that allows the doctor or therapist to design a specific path and size of the road for different patients according to their needs.

The doctor or therapist can check on a chart, the results of the games of each patient or all patients, or state a specific level of game, and it is also possible to see on a graph, an specific exercise that the patient performed during a period of time.

Frog game.

Frog game is for exercising upper limbs movements and also exercising thumb movements (press the button to turn on the light).

This game works in the following way: the frog has to eat all cakes by pointing the marker in each cake and press the button, turning on the light to eat the cake, the frog must avoid hitting the bombs, when it collides with a bomb, the game emits error sound and it counts an error; when all cakes are eaten, the game emits a clapping sound meaning that user has finished the level of the game. Either user or therapist can choose the difficulty of the game in the same level, eating cakes in order or not, with bombs or not, bombs with movements or not.

The doctor and therapist can also create their own levels of this game, selecting number of bombs and cakes, marking their positions, selecting the direction of the routes of each bomb. This game also offers the search of the results of the patient's performance, using a chart or graph. Such as shown in Figure 5.

Fig. 5. Graph result

4 Study case

Our study case has been done during 1 month to 6 months with children aged 6-10 years with cerebral palsy. Table 1 shows the observation and progress of these children. We have also tried during a period of time with children from 1 and half to 4 years with cerebral palsy, but this game is difficult for them, they get distracted very easily, most of them cannot play this game alone, they need their parent's help. So we just didn't use these cases.

During the test time with children from 6 to 10 years, they made very positive progress. One of them can do the exercises alone, and every time he makes better exercises, points more exactly the road with the marker and finishes with less time and less errors. In the beginning he couldn't reach higher levels, after a period of therapy he could reach a higher level, now he can finish high levels within a short time and having few or cero mistakes. Table 2 shows the progress comparing the average results of beginning and after 4 month.

Table 1. Patient's data

	P1	P2	P3	P4	P5
Age	10 years	9 years	8 years	6 years	7 years
Exercised hand	Left	Right	Left	Right	Right
Time of exercise	6 month	1.5 month	2 month	1 month	2 month
Time of each therapy	30 minutes	30 minutes	30 minutes	30 minutes	30 minutes
Position of	standing	sitting	sitting	standing	standing
Hand common use	Right	Right	Left	Right	Right
Description at beginning stage	Only arrive level 4	Cannot hold pen well and can't paint with strength	Cannot do exercise alone	Cannot do exercise alone	Cannot do exercise alone
progressing	Arrive to all levels with good results, can do fine work like opening the door with key, catch a ball	Can cut and paste on a straight line, paint with more strength and inside of frame	Can exercise until level 3	Can exercise until level 5	Can exercise until level 4

The case of frog game has also had good progress, child P1 has participated for 3 months and a half. In the beginning P1 couldn't press the button with his thumb of the left hand, which is the hand he's been doing the therapies with. He couldn't even holding the marker well, he needed right hand support, after 3 months the child can hold marker with left hand very well and press the button with his thumb.

For children with cerebral palsy, the present project is not only an exercise therapy to upper limbs but it is also an exercise to identify colors, objects and through the different sounds they can identify its meaning. In one single therapy several goals are achieved; both rehabilitation of upper limb, also color recognition and identification of objects and sounds. This way children find interesting to do their rehabilitation therapy.

Table 2. Comparing average result of child P1

Level	In the beginning		4 month later	
	Time (s)	error	Time (s)	error
1	26.516	3.33	19.100	0.33
1 A	16.171	4.33	13.235	2
2	9.047	1	11.67	0.66
3 A	53.500	13.33	28.846	1
4	24.166	3	12.046	0.33
4 A	21.400	4	15.520	0.66
5 A	99.635	61.5	38.604	3
5 B	46.133	2.25	12.197	0
6 A	47.154	7.75	32.045	1.667
6 B	69.639	19	20.460	3
6 C	93.631	5	31.368	3.66

5 Conclusion

The present paper shows the use of image processing in the rehabilitation area, offering neuromuscular rehabilitation games to help patients, and showing through the case study very positive results and progress of children who participated in this case study. This gives us more confidence that the rehabilitation games are very good tools to help in rehabilitation therapies and that they are not only for the therapies of movement but also offer another extra advantages. The best thing of this project is that it can be used for different king of rehabilitations, different therapies, to satisfying needs of different patients, only by changing the game.

The next step of this project is to continue creating different types of games for different therapies in order to offer a software tool containing various games for diverse needs of patient.

References

1. David Furió, Santiago González-Ganceda, M-Carmen Juan, Ignacio Seguí, Noemí Rando: Evaluation of learning outcomes using an educational iPhone game vs. traditional game. In: Computers & Education vol. 64 (May 2013) 1-23
2. Remus-Catalin Prodan, Stefan-Gheorghe Pentiuc, Radu-Daniel Vatavu: An efficcient solution for hand gesture recognition from video sequence. In: Advances in Electrical and Computer Engineering, Vol. 12 No. 3 (2012)
3. Wei Peng, Julia C. Crouse, Jih-Hsuan Lin: Using active video games for physical activity promotion, a systematic reveiw of the current state of research. In: Health education & behavior vol. 40 no. 2 (april 2013)
4. Cristina Stolojescu-CriŞan, Ştefan Holban: A comparison of X-Ray image segmentation techniques. In: Advances in Electrical and Computer Engineering Engineering Volume 13, number 3 (2013)

5. S. Patel, H. Park, P. Bonato, L. Chan and Mary odgers: A review of wearable sensors and systems with application in rehabilitation. In: Journal of NeuroEngineering and Rehabilitation (2012)
6. Andrew Patrick Hayes W., Nick Preston, Raymond Holt, Matthew Allsop,Martin Levesley, Bipinchandra Bhakta: Engaging children in healthcare technology design: developing rehabilitation technology for children with cerebral palsy. In: Journal of Engineering Design Vol. 21, No. 5 (October 2010) 579–600
7. Marcus King, Juha M. Hijmans, Michael Sampson, Jessica Satherley, Leigh Hale: Home-based stroke rehabilitation using computer gaming. In: New Zealand Jounal of Physiotherapy (2012)
8. Kate Laver, Stacey George, Julie Ratcliffe, Maria Crotty: Virtual reality stroke rehabilittion - hype or hope? In: Australian Occupational Therapy Journal vol. 58 (2011) 215-219
9. Pilar Póo Argüelles: Parálisis cerebral infantil. In: Protocolos Diagnóstico Terapeúticos de la AEP: Neuología Pediátrica (2008)
10. José Ramón Corderí: Rehabilitacion en enfermedades neuromusculares. In: http://www.adm.org.ar/wp-content/uploads/2012/09/terapia_fisica.pdf (Febrero 2013)

A robust data association method based on B-Splines for SPLAM in complex environments

Alfredo Toriz P.[‡], Abraham Sánchez L., Rogelio Gonzalez V. and Beatriz Bernabe L.

[‡]Universidad Popular Autónoma del Estado de Puebla, Pue-México
Benemérita Universidad Autónoma de Puebla, Pue-México
alfredo.toriz@upaep.mx, asanchez@cs.buap.mx,rgonzalez@cs.buap.mx, bety@cs.buap.mx

Abstract. In this work, we present a robust data association method for solving the SPLAM (Simultaneous Planning Localization and Mapping) problem in complex environments. The use of B-Splines curves as representation tool is one of the most important contributions. The problem of data association for SPLAM was initially addressed by comparing the control points forming the curves that represent the observed obstacles at a given moment, to the obstacles stored in the map, correlating those who are close enough. Then, the curvature of the related B-spline is obtained to extract characteristic points (inflection points and corners) contained in the curves. Finally, the matching information was used to correct the position of the robot and the detected obstacles. We carried out numerous experiments by using real and simulated information in order to validate the processes and algorithms proposed in our approach. Our method achieves a great precision in the map construction of complex environments, which is nearly impossible with techniques that currently exist.

1 Introduction

One of the fundamental challenges in robotics is to obtain robust and efficient mechanisms to model increasingly complex environments by using mobile robots or other kinds of robots. Many approaches have dealt with this problem without great success, since most of these methods try to extract geometric features and represent their positions on the map [1], or to discretize the space into cells [2] and then classify each as occupied or empty. These approaches are all limited to well-structured environments of limited size.

From these limitations, new methods have emerged in the literature that try to deal with this problem. One recent work uses B-Spline curves as a representation form for describing complex geometries with obstacles [3]. However, this new form of representation itself presents a new challenge in identifying similar environmental characteristics from successive scans and moments, so errors in the robot odometry can be corrected as the map of the environment is built. The approach presented by Pedraza et al., solves the problem of the representation

©L. Berthet, J. C. Chimal, E. A. Santos, E. Castillo
Advances in Computer Science
Research in computing Science 81, 2014 pp. 13-25

of complex geometries with obstacles by using a simple method of data association. This method involves projecting rays from a laser sensor from observed objects in the estimated odometry position to the located obstacles contained in the map, regardless of whether these points are really consistent. While this is represents a real advance in the field, the method may lead to inconsistencies in the created map.

In this way, data association is undoubtedly the most critical aspect of the algorithm for map construction and it is considered by some authors to be the most difficult problem in this field of robotics. Therefore, the correct association between the sensed characteristics during an observation and the generated map features are essential for a consistent map construction, since any false association could invalidate the whole process.

The process of data association has received great attention in the robotics literature in recent years [4], [5] and many solutions have been proposed; However, most of these approaches are limited to simple features such as lines and points (preventing their use in complex environments), simple geometries, or simple curves where the association algorithms are very rudimentary. Although these techniques are appropriate and allow a full SPLAM in some cases, it is still necessary to develop and implement new strategies that may be applicable to more general cases without specific geometries.

In this paper, we present a new data association method for map construction of complex environments based on B-Spline curves, which are used to represent the obstacles in the map. Thus, our approach proposes an improvement on current methods of data association by fully exploiting the information contained in the parametric description of arbitrary geometries. With the creation and validation of an advanced map modeling method such as B-Spline curves, the properties of this representation (curves, length of curves, inflection points) are used to obtain a robust data association method applicable to most, if not all, mapping situations.

2 Data association based on B-Splines

2.1 Data management

When the robot gets a new set of measures of its environment through its perception system, the surrounding world is just a set of points $p_i \in \Re^2$ linked to each other only by the logic provided by the scanning sensor (Figure 1). With these raw data, the first objective is to clearly identify the object associated with the measurements grouped into subsets to finally get the B-spline curves that represent the portions of the detected objects as closely as possible.

The segmentation of the raw data was performed with the adaptive cluster method proposed by Dietmayer [6] with the goal of obtaining measurements subgroups belonging to different objects. Once the subsets or clusters were obtained, the measurements corresponding to each detected object were approximated us-

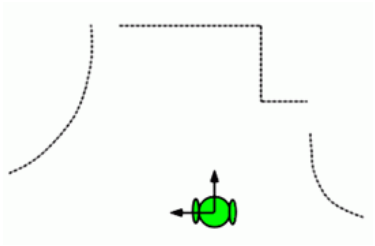


Fig. 1. Measurement points in space $p_i \in \Re^2$ relative to an origin (the scanning robot) obtained with a laser scanning process.

ing unclamped B-spline curves of degree 3 as shown in the next equation:

$$D_k = C(t_k) = \sum_{i=0}^{n} N_{i,p}(t_k) X_i \text{ for } 0 \leq k \leq m. \tag{1}$$

where $\sum N_p(u)X$ is the Cox-de-Boor recursion formula to obtain a B-Spline curve, and the parameter t_k represents a point on the curve [7].

Although the representation based on B-Splines significantly reduced the noise caused by the measurement errors of the system; a smoothing of the curve was performed using a Gaussian filter to ensure that the process did not find false characteristic points. Then, an evolved version D_σ of the curve D was processed:

$$D_\sigma = \{x(u,\sigma), y(u,\sigma)\} \tag{2}$$

where

$$x(u,\sigma) = x(u) \otimes g(u,\sigma),\ y(u,\sigma) = y(u) \otimes g(u,\sigma) \tag{3}$$

Here, $\otimes$ represents the convolution operator and $g(u,\sigma)$ denotes a Gaussian filter of amplitude σ, chosen conveniently and used only to remove the noise. While the valuable information about the curve was preserved, this selection process made the value to be very low. Since our proposed method normally works with open curves, a certain number of proportional points to the double of the full width at half-maximal (FWHM) of a Gaussian, were symmetrically offset on both ends of the curve when smoothing.

Finally, in order for the association process to effectively execute, it was necessary to ensure the invariance in the resolution of the curve. so, each discrete B-Spline curve was therefore stored taking equidistant points with a distance ε between each point:

$$dist(\sum N_p(u)X, \sum N_p(u+1)X) \cong \varepsilon \tag{4}$$

Thus, a parametric vector containing the B-Spline was obtained. In addition to the described process, a restriction on the length of the curve was applied; this due to the fact that too-small objects do not provide sufficient information and therefore it is not interesting to include them on the map. Also, although

our method is currently designed for use in static environments, this restriction allows filtering dynamic elements (people for example) that are not meant to be included in the map. Once the B-Splines were obtained and chosen, specific features contained in the curves were identified that were of great importance in the localization process. Essentially, two types of features were sought in the curves: inflection points and corners.

The corner type is very common feature representing a sharp point on a curve that is generally differentiable. On the other hand, the concept of inflection points in a very general way refers to the point on the curve where this passes from concave to convex or viceversa.

The process for obtaining both characteristics in the curve is based on the curvature scale space (CSS) [8] which is used to retrieve invariant geometric characteristics.

2.2 Search of characteristics in B-Splines curves

The search for particular characteristics in the curves that define objects is completely based on the curvature of the B-Spline, which is defined as the local measure that indicates how much a curve has moved away from a straight line. More formally, the curvature of a point $X_u = [x_u, y_u]$ on the B-Spline is defined as the amount equal to the inverse of the radius of the osculating circle at the point (the circle that touches tangentially to the curve in the X_u point), which means that the smaller is the radius of this circle ρ largest is the curvature at this point $(1/\rho)$. The formula for processing the curvature can be expressed as:

$$k(u,\sigma) = \frac{\dot{x}(u,\sigma)\ddot{y}(u,\sigma) - \ddot{x}(u,\sigma)\dot{y}(u,\sigma)}{(\dot{x}(u,\sigma)^2 - \dot{y}(u,\sigma)^2)^{(3/2)}} \tag{5}$$

Where, in accordance with the convolution properties, the derivatives of each element could be easily calculated since the exact forms of the first and second derivatives of the used Gaussian kernels $\dot{g}(u,\sigma)$ and $\ddot{g}(u,\sigma)$ are known. Thus,

$$\dot{x}(u,\sigma) = \frac{\partial}{\partial u}(x(u) \otimes g(u,\sigma)) = x(u) \otimes \dot{g}(u,\sigma) \tag{6}$$

$$\ddot{x}(u,\sigma) = \frac{\partial^2}{\partial^2 u}(x(u) \otimes g(u,\sigma)) = x(u) \otimes \ddot{g}(u,\sigma) \tag{7}$$

$$\dot{y}(u,\sigma) = y(u) \otimes \dot{g}(u,\sigma) \tag{8}$$

$$\ddot{y}(u,\sigma) = y(u) \otimes \ddot{g}(u,\sigma) \tag{9}$$

Considering the above, the list of curvatures of the B-Spline is obtained by using the equation 5, i.e., the obtaining of the inflection points in the curve is carried out simply just searching for a change in the sign between two consecutive curvatures. Moreover, the searching of corners is based on the work of He et al., presented in [9]. Here, the detector is based on local and global properties of the curvature (Figure 2).

Since our proposed method uses open curves in most cases, we considered only corners delimited on both sides by inflection points as a safety and certainty method.

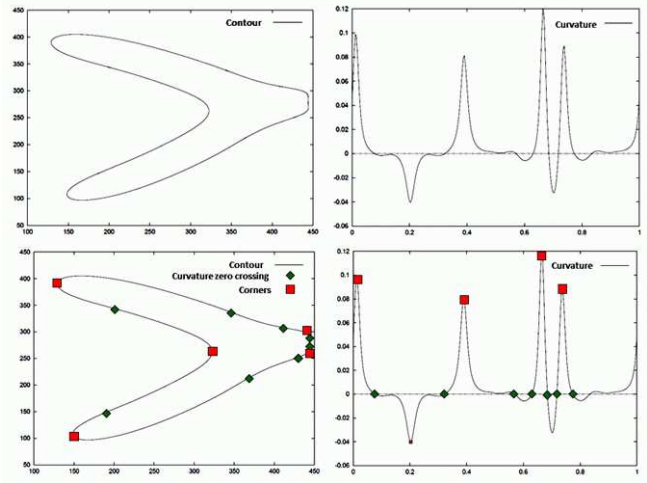


Fig. 2. Curvature zero crossings (inflections) and corners detection shown with some example curves to validate the detection and classification processes.

2.3 Data association

In order to understand the proposed approach, this section presents an illustrative example, which will be described in each stage of development. Figure 3 shows an example that corresponds to the evolution of the robot exploration until the instant $k + s$, where three objects $B_{O,1}$, $B_{O,2}$ and $B_{O,3}$ are detected (shown in red) in the sensor's range of vision (circumference in dotted line). This image highlights the odometric position as an a priori approximation of the robot's true position, obtained from the robot's motion model since errors and the sensory noise will lead the robot to a different real position and not to the expected one. For this reason, the objects $B_{O,1}$, $B_{O,2}$ and $B_{O,3}$ appear displaced and do not overlap exactly with the objects to which they belong.

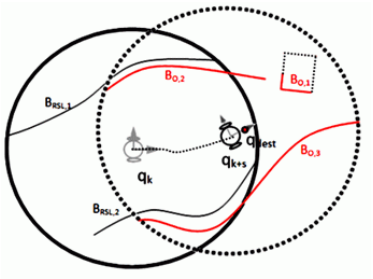


Fig. 3. The robot at the odometric position q_{k+s} with three obstacles detected within its detection range.

The B-Splines association will consider initially only the points of the control polygon that generated the curves. In this step, the distances between the points of the control polygons of the all objects (those contained in the Local Safe Region (see [10] for more details about the exploration method) and those observed in the position q_{k+s}) are obtained by associating the observed curves with the reference curves using the following criteria.

$$\min(dist(X_{LSR,i}, X_{O,j})) \leq d_{min}, \ \{ \begin{matrix} i = 1, ..., n_m \\ j = 1, ..., n_0 \end{matrix} \tag{10}$$

Where $X_{O,j}$ and $X_{LSR,i}$ are the control points of the observed splines and the splines in the current LSR respectively, $dist(X_{LSR,i}, X_{O,j})$ represents the Euclidean distance between two points, and n_m and n_0 are the number of control points of the splines in the map and of the observed splines, respectively. At the end of this first stage, the splines with a minimum number μ_{min} of related control points will be associated.

In Figure 4, the curves $(B_{LSR,1}, B_{O,2})$ and $(B_{LSR,2}, B_{O,3})$ have been associated, since they have 5 and 7 related points of their control polygon respectively, which are obtained with equation 10.

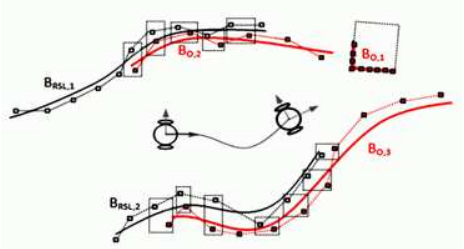


Fig. 4. Rough association computed with the control points of the curves.

If the curves are related at this point, the next step is to perform a fine association; for which, corners and inflection points contained in the related curves will be searched. If successful, the items found will be used for performing an accurate association between each pair of curves $(B_{LSR,i}, B_{O,j})$. The information about the type of characteristic and the curvature at that point will be used to avoid errors in this association step (Figure 5). If in some related curves no associated elements were found (lines or curves too smooth, for example), the fine association process will be executed to finding the closest points to the ends of the related curves as the authors proposed in [3].

Once all the elements are related, a search of the initial and final points of the related curves is performed. Taking the inflection points or the more extreme corners as starting points contained in both related curves, one can use some continuous points on the parametric curve to its end. The number of considered points is the maximum number of elements that can be taken in the curve

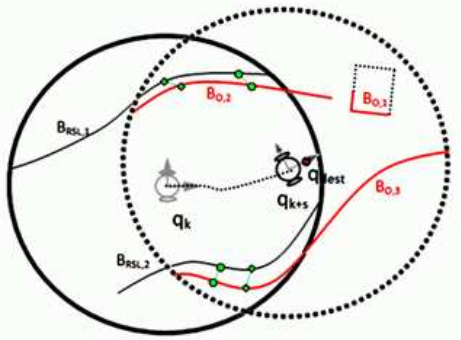


Fig. 5. Association of the inflection points between the curves and the corners of the local safe region and the observed curves. Green circles represent inflection points while diamonds represent corners.

segment of the smaller length of the two related points from the most extreme characteristic point toward the ends of the curve. This can be seen in detail in the related curves ($B_{LSR,2}, B_{O,3}$) in Figure 6. Here, one can note that the initial point, represented by the blue circle, was taken by choosing six elements of the parametric curve (they are showed in blue dashed line) from the inflection point, represented by the green circle, to the end of the curve. The length of the curve segment $B_{O,3}$ from the beginning of the curve to the inflection point, it is larger than the curve segment $B_{LSR,2}$ from the beginning of the curve to the inflection point and therefore the elements of the curve segment with shorter length will surely be contained in the another curve segment of greater length.

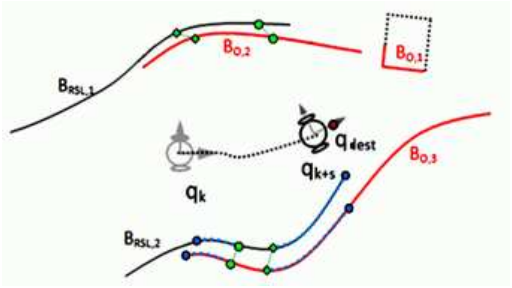


Fig. 6. Example of how the initial and the end point of the related curve segments are found.

The same process is performed at the end of the curve, where 15 points of the parametric curve were taken from the corner, represented with green diamond. The curve segment of smaller length from the corner at the end of the curve belonging to $B_{LSR,2}$contains only 15 points. Finally, when the association

process ended, the related curves will have an appearance similar to those shown in Figure 7.

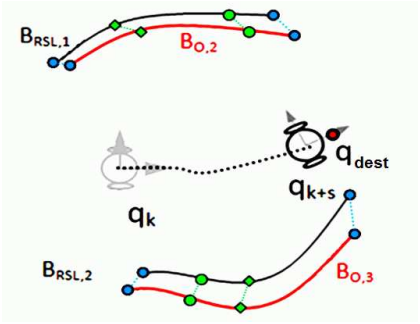


Fig. 7. The curve segments related to the described process of data association. Blue circles represent the ends of related curves, while green circles and green diamonds represent inflection points and the corners of the curve respectively.

3 Experimental results

Although the data association is a general problem that is applicable to any SPLAM algorithm, in this section we decided to implement our proposal by using the Extended Kalman Filter for SPLAM. This was done in order to make a comparison of our proposal with the proposal submitted by Pedraza et al. [3], since this is one of the few works that has grappled with the problem of complex environment representation using the same description of obstacles used in this work, though with a different approach to data association for our work. We have conducted numerous experiments with both real and simulated data in order to verify and validate the approach proposed in this work. The data obtained with simulation experiments have allowed the verification of the accuracy and consistency properties of the algorithm by comparisons with other existing and validated algorithms. Additionally, experiments with different data from real environments were used to verify the applicability and effectiveness of these techniques.

A simulated robot and the real Pioneer P3DX robot equipped with front and rear bumper arrays, a ring of eight forward ultrasonic transducer sensors (range-finding sonar) and a Hokuyo URG-04Lx laser range finder (with a detection range from 0.02 to 4 meters with a standard deviation (σ_L) of 1%, an angular resolution of 0.36 degrees and a scan angle of 240 degrees) were used in the experiments. The ultrasonic transducer sensors were used to obtain environmental information in the 120 degree range, since the laser range finder cannot see at this angle. The LIRMM laboratory environment was used in the experimental and simulation tests (the environment had several corridors), see Figure 8.

The data association strategy developed in this work and any other solution proposed in this field could be validated using a computational performance criterion, map quality and consistency of the algorithms. However, unlike the methods whose environment representation is based on specific geometries and where much of the information acquired by the sensors is wasted, our approach exploits the maximum amount of possible information provided by the sensors avoiding dangerous simplifications.

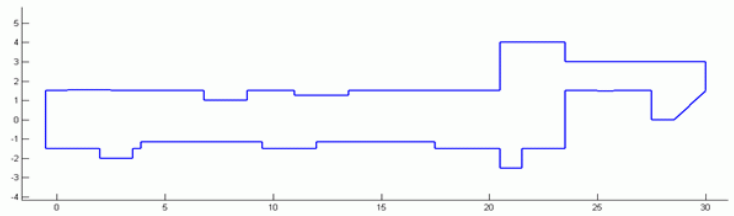


Fig. 8. The LIRMM environment.

Although the environment used in our experiments does not represent a serious challenge to the traditional algorithms, the use of a data association method based on B-Spline curves allows linking segments of continuous data, without the need for segmentation of information into smaller pieces. From the above, we find that the data association algorithm based on B-Splines, extends the capabilities of the SPLAM algorithms to more general situations. Figure 9 shows the results of SPLAM experiments we have developed as part of our research.

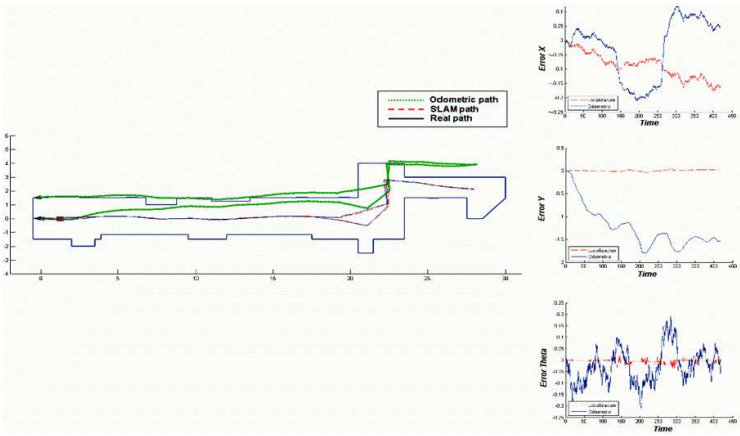


Fig. 9. Accuracy and consistency experiments for the EKF-SPLAM based on the data association approach proposed in this work.

Figure 9 shows the real trajectory of the robot (blue solid line), the odometric trajectory (green dotted line) and the trajectory obtained by the EKF-SPLAM method (dotted red line). On the other hand, when there is certainty about the real trajectory of the robot (as in the simulations) is possible to perform some checks to give an idea of the quality of the algorithms from the point of view of consistency. For this reason, it has been possible to include in the representation of the odometric error in X, Y and θ on the right part of the figure (blue line), as well as the localization errors (red dotted line). From the data obtained with our approach and taking as reference the errors shown by the SPLAM method presented by Pedraza et al. [3], we conclude that the proposed method of data association notably improve the quality of the map and the errors as shown in Figure 10.

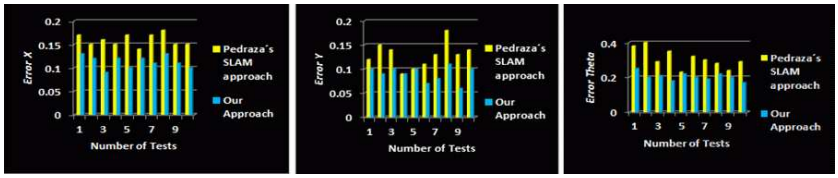


Fig. 10. Errors obtained with SPLAM strategies, for the simple data association [3] and for the proposed data association in this work.

Finally, in this section we present experiments with real data in real environments to validate the results presented in the simulation step. In all tests, the maps were obtained by considering only the odometric information reported by the robot and the maps obtained after applying our SPLAM implementation with our proposed data association. Figure 11 shows the real environment. The map obtained with our SPLAM strategy is shown in Figure 12a, while the odometric map obtained using only odometry information is shown in Figure 12b.

Fig. 11. The real environment used for the tests.

Considering that our approach has been proposed to build maps of complex environments, another experiment was performed using the real environment

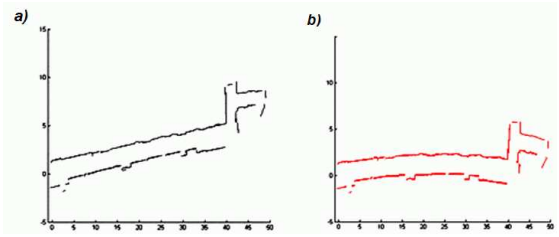


Fig. 12. a) Map built with the proposed SPLAM strategy. b) Map built with only odometry information.

shown in Figure 13. In this example, we can see that the construction of the map would be impossible using a conventional method based on lines and points. Figure 14 shows the map obtained with our approach.

Fig. 13. LIRMM environment with objects.

4 Conclusions

We have proposed a method for data association based on the curvatures analysis of the related B-Splines curves, for solving the SPLAM (Simultaneous Planning Localization and Mapping) problem in complex environments. Here, we use the techniques of digital images CSS [8], used in the pattern recognition field, inflection points and corners extraction among others. This association mechanism not only established a robust correspondence between the observations made by the robot and the objects contained in the map, but it facilitated parametric correspondence between each pair of associated representative elements.

After testing the proposed method in both simulated and real complex environments, and comparing the performance of the system to standard algorithms like EKF_SPLAM, we conclude that this functionality was well demonstrated and supports the strength of this new approach to this family of mapping problems.

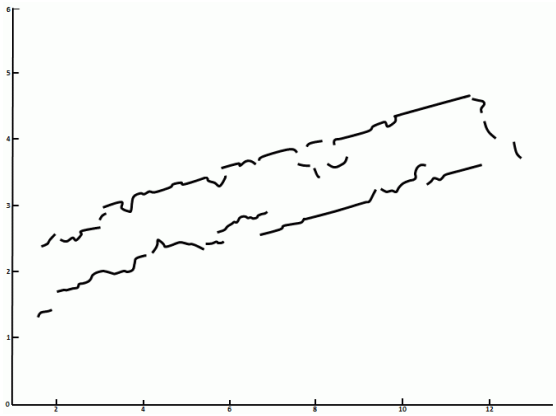


Fig. 14. Final map of the complex environment.

References

1. S. Zhang, L. Xie and M. Adams, "Gradient model based feature extraction for simultaneous localization and mapping in outdoor applications", *IEEE Control, Automation, Robotics and Vision Conference*, (2004) 431-436
2. G. Grisetti, C. Stachniss and W. Burgard, "Improving Grid-based SLAM with Rao-Blackwellized Particle Filters by Adaptive Proposals and Selective Resampling", *IEEE International conference on Robotics and Automation*, (2005) 2432-2437
3. L. Pedraza, D. Rodriguez-Losada, F. Matia, G. Dissanayake and J. V. Miro, "Extending the Limits of Feature-Based SLAM with B-Splines", *IEEE Transactions on Robotics*, Vol. 25, Issue 2, (2009) 353-366
4. A. J. Cooper, "A comparison of data association techniques for Simultaneous Localization and Mapping", *PhD Thesis, Massachusetts Institute of Technology - Dept. of Aeronautics and Astronautics*, (2005)
5. S. Thrun, M. Montemerlo, D. Koller, B. Wegbreit, J. Nieto and E. Nebot, "FastSLAM: An efficient solution to the simultaneous localization and mapping problem with unknown data association", *Journal of Machine Learning Research*, (2004)
6. K. C. J Dietmayer, J. Sparbert and D. Streller, "Model based object classification and object tracking in traffic scenes from range images", *Proc. of the IV IEEE Intelligent Vehicles Symposium*, (2001)
7. C. de Boor, "A practical guide to Splines", *Springer-Verlag, New York*, (1978)
8. F. Mokhtarian, "Silhouette-based isolated object recognition through curvature scale space", *IEEE Pattern Analysis and Machine Intelligence*, Vol. 17, No. 5, (1995) 539-544
9. X. C. He and N. H. C. Yung, "Corner detector based on global and local curvature properties", *Optical Engineering*, 47(5) (2008)
10. A. Toriz P., A. Sánchez L., R. Zapata and M. A. Osorio L., "Mobile robot SPLAM for robust navigation", *Journal Research in Computing Science*, Vol. 54, (2011) 295-306

Image Quantizer based on Contrast Band-Pass Filtering

Jaime Moreno[1,2], Oswaldo Morales[1], and Ricardo Tejeida[1]

[1] National Polytechnic Institute of Mexico,
IPN Avenue, Lindavista, Mexico City, 07738, Mexico.
[2] XLIM Laboratory, Signal, Image and Communications Department,
University of Poitiers, 86962 Futuroscope, France.
jmorenoe@ipn.mx

Abstract The aim of this work is to explain how to apply perceptual criteria in order to define a perceptual forward and inverse quantizer. We present its application to the H*i*-SET coder. Our approach consists in quantizing wavelet transform coefficients using some of the human visual system behavior properties. Taking in to account that noise is fatal to image compression performance, because it can be both annoying for the observer and consumes excessive bandwidth when the imagery is transmitted. Perceptual quantization reduces unperceivable details and thus improve both visual impression and transmission properties. The comparison between JPEG2000 coder and the combination of H*i*-SET with the proposed perceptual quantizer (χSET) shows that the latter is not favorable in PSNR than the former, but the recovered image is more compressed (less bit-rate) at the same or even better visual quality measured with well-know image quality metrics, such as MSSIM, UQI or VIF, for instance.

Keywords: Human Visual System, Contrast Sensitivity Function, Perceived Images, Wavelet Transform, Peak Signal-to-Noise Ratio, No-Reference Image Quality Assessment, JPEG2000.

1 Introduction

Digital image compression has been a research topic for many years and a number of image compression standards has been created for different applications. The JPEG2000 is intended to provide rate-distortion and subjective image quality performance superior to existing standards, as well as to supply functionality [1]. However, JPEG2000 does not provide the most relevant characteristics of the human visual system, since for removing information in order to compress the image mainly information theory criteria are applied. This information removal introduces artifacts to the image that are visible at high compression rates, because of many pixels with high perceptual significance have been discarded.

Hence, it is necessary an advanced model that removes information according to perceptual criteria, preserving the pixels with high perceptual relevance regardless of the numerical information. The Chromatic Induction Wavelet Model

©L. Berthet, J. C. Chimal, E. A. Santos, E. Castillo
Advances in Computer Science
Research in computing Science 81, 2014 pp. 25-35

presents some perceptual concepts that can be suitable for it. Both CBPF and JPEG2000 use wavelet transform. CBPF uses it in order to generate an approximation to how every pixel is perceived from a certain distance taking into account the value of its neighboring pixels. By contrast, JPEG2000 applies a perceptual criteria for all coefficients in a certain spatial frequency independently of the values of its surrounding ones. In other words, JPEG2000 performs a global transformation of wavelet coefficients, while CBPF performs a local one.

CBPF attenuates the details that the human visual system is not able to perceive, enhances those that are perceptually relevant and produces an approximation of the image that the brain visual cortex perceives. At long distances the lack of information does not produce the well-known compression artifacts, rather it is presented as a softened version, where the details with high perceptual value remain (for example, some edges).

2 JPEG2000 Global Visual Frequency Weighting

In JPEG2000, only one set of weights is chosen and applied to wavelet coefficients according to a particular viewing condition (100, 200 or 400 dpi's) with fixed visual weighting[1, Annex J.8]. This viewing condition may be truncated depending on the stages of embedding, in other words at low bit rates, the quality of the compressed image is poor and the detailed features of the image are not available since at a relatively large distance the low frequencies are perceptually more important. The table 1 specifies a set of weights which was designed for the luminance component based on the CSF value at the mid-frequency of each spatial frequency. The viewing distance is supposed to be 4000 pixels, corresponding to 10 inches for 400 dpi print or display. The weight for LL is not included in the table, because it is always 1. Levels 1, 2, ... , 5 denote the spatial frequency levels in low to high frequency order with three spatial orientations, *horizontal, vertical* and *diagonal*.

Table 1. Recommended JPEG2000 frequency (s) weighting for 400 dpi's ($s = 1$ is the lowest frequency wavelet plane).

s	*horizontal*	*vertical*	*diagonal*
1	1	1	1
2	1	1	0.731 668
3	0.564 344	0.564 344	0.285 968
4	0.179 609	0.179 609	0.043 903
5	0.014 774	0.014 774	0.000 573

3 Perceptual Forward Quantization

3.1 Methodology

Quantization is the only cause that introduces distortion into a compression process. Since each transform sample at the perceptual image $\mathcal{I}_\rho$ is mapped independently to a corresponding step size either Δ_s or Δ_n, thus $\mathcal{I}_\rho$ is associated with a specific interval on the real line. Then, the perceptually quantized coefficients $\mathcal{Q}$, from a known viewing distance d, are calculated as follows:

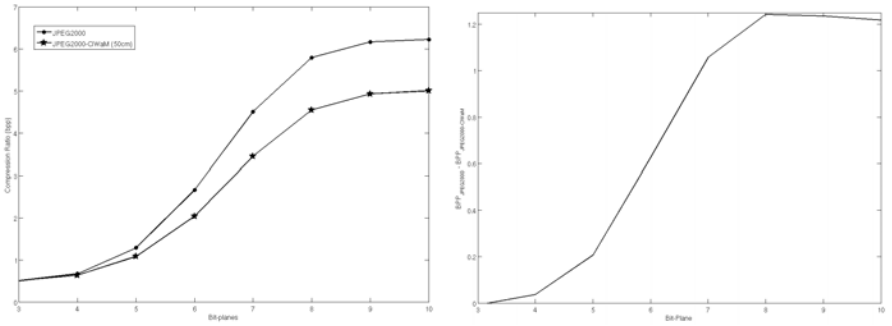
$$\mathcal{Q} = \sum_{s=1}^{n} \sum_{o=v,h,d} sign(\omega_{s,o}) \left\lfloor \frac{|\alpha(\nu, r) \cdot \omega_{s,o}|}{\Delta_s} \right\rfloor + \left\lfloor \frac{c_n}{\Delta_n} \right\rfloor \tag{1}$$

Unlike the classical techniques of Visual Frequency Weighting (VFW) on JPEG2000, which apply one CSF weight per sub-band [1, Annex J.8], Perceptual Quantization using CBPF (ρSQ) applies one CSF weight per coefficient over all wavelet planes $\omega_{s,o}$. In this section we only explain Forward Perceptual Quantization using CBPF (F-ρSQ). Thus, Equation 1 introduces the perceptual criteria of Perceptual Images to each quantized coefficient of Equation of Dead-zone Scalar Quantizer. A normalized quantization step size $\Delta = 1/128$ is used, namely the range between the minimal and maximal values at $\mathcal{I}_\rho$ is divided into 128 intervals. Finally, the perceptually quantized coefficients are entropy coded, before forming the output code stream or bitstream.

3.2 Experimental Results applied to JPEG2000

The Perceptual quantizer F-ρSQ in JPEG2000 is tested on all the color images of the *Miscellaneous volume* of the University of Southern California Image Data Base[2]. The data sets are eight 256×256 pixel images and eight 512×512 pixel images, but only visual results of the well-known images *Lena*, *F-16* and *Baboon* are depicted, which are 24-bit color images and 512×512 of resolution. The CBPF model is performed for a 19 inch monitor with 1280 pixels of horizontal resolution at 50 centimeters of viewing distance. The software used to obtain a JPEG2000 compression for the experiment is *JJ2000*[3]. Figure 1(a) shows the assessment results of the average performance of color image compression for each bit-plane using a Dead-zone Uniform Scalar Quantizer (SQ, function with heavy dots), and it also depicts the results obtained when applying F-ρSQ(function with heavy stars). Using CBPF as a method of forward quantization, achieves better compression ratios than SQ with the same threshold, obtaining better results at the highest bit-planes, since CBPF reduces unperceivable features. Figure 1(b) shows the contribution of F-ρSQ in the JPEG2000 compression ratio, for example, at the eighth bit-plane, CBPF reduces 1.2423 bits per pixel than the bit rate obtained by SQ, namely in a 512×512 pixel color image, CBPF estimates that 39.75KB of information is perceptually irrelevant at 50 centimeters.

Figure 2 depicts examples of recovered images compressed at 0.9 bits per pixel by means of JPEG2000 (a) without and (b) with F-ρSQ. Also these figures



(a) (b)

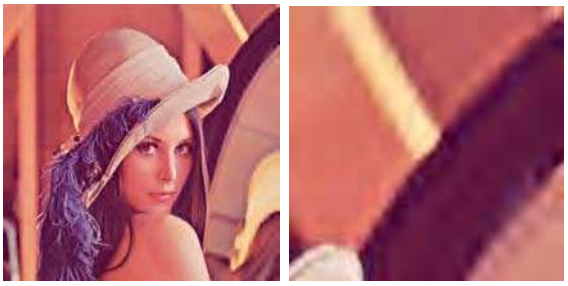
Figure 1. (a)JPEG2000 Compression ratio (bpp) as a function of Bit-plane. Function with heavy dots shows JPEG2000 only quantized by the dead-zone uniform scalar manner. While function with heavy stars shows JPEG2000 perceptually pre-quantized by F-ρSQ. (b)The bit-rate decrease by each Bit-plane after applying F-ρSQ on the JPEG2000 compression.

show that the perceptual quality of images forward quantized by ρSQ is better than the objective one. Also, figure 3 shows examples of recovered images of *Baboon* compressed at 0.59, and 0.45 bits per pixel by means of JPEG2000 (a) without and (b) with F-ρSQ. In Fig. 3(a) PSNR=26.18 dB and in Fig. 3(b) PSNR=26.15 dB but a perceptual metrics like WSNR [4], for example, assesses that it is equal to 34.08 dB. Therefore, the recovered image Forward quantized by ρSQ is perceptually better than the one only quantized by a SQ. Since the latter produces more compression artifacts, the ρSQ result at 0.45 bpp (Fig. 3(b)) contains less artifacts than SQ at 0.59 bpp. For example the *Baboon*'s eye is softer and better defined using F-ρSQ and it additionally saves 4.48 KB of information.

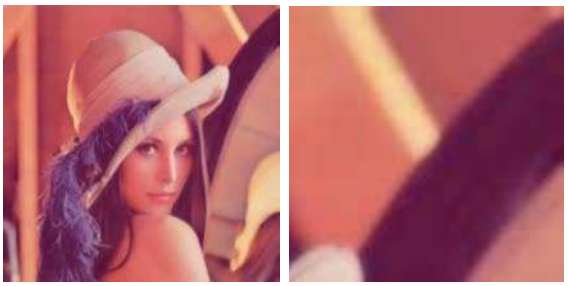
4 Perceptual Inverse Quantization

The proposed Perceptual Quantization is a generalized method, which can be applied to wavelet-transform-based image compression algorithms such as EZW, SPIHT, SPECK or JPEG2000. In this work, we introduce both forward (F-ρSQ) and inverse perceptual quantization (I-ρSQ) into the H*i*-SET coder. This process is shown in the green blocks of Fig. 4. An advantage of introducing ρSQ is to maintain the embedded features not only of H*i*-SET algorithm but also of any wavelet-based image coder. Thus, we call CBPF Perceptual Quantization + H*i*-SET = cH*i*-SET or χSET.

Both JPEG2000 and χSET choose their VFWs according to a final viewing condition. When JPEG2000 modifies the quantization step size with a certain visual weight, it needs to explicitly specify the quantizer, which is not very suitable for embedded coding. While χSET neither needs to store the visual



(a) JPEG2000 PSNR=31.19 dB.



(b) JPEG2000-F-ρSQ PSNR=27.57 dB.

Figure 2. Examples of recovered images of Lenna compressed at 0.9 bpp.

weights nor to necessarily specify a quantizer in order to keep its embedded coding properties.

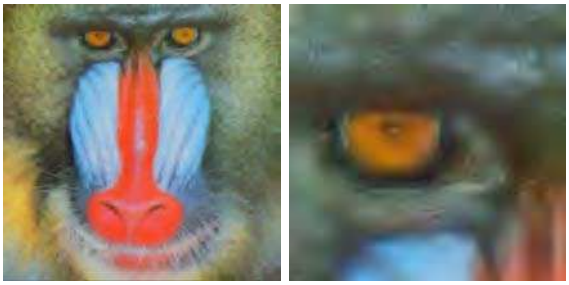
The main challenge underlies in to recover not only a good approximation of coefficients $\mathcal{Q}$ but also the visual weight $\alpha(\nu, r)$(Eq. 1) that weighted them. A recovered approximation $\widehat{\mathcal{Q}}$ with a certain distortion Λ is decoded from the bitstream by the entropy decoding process. The VFWs were not encoded during the entropy encoding process, since it would increase the amount of stored data. A possible solution is to embed these weights $\alpha(\nu, r)$ into $\widehat{\mathcal{Q}}$. Thus, our goal is to recover the $\alpha(\nu, r)$ weights only using the information from the bitstream, namely from the Forward quantized coefficients $\widehat{\mathcal{Q}}$.

Therefore, our hypothesis is that an approximation $\widehat{\alpha}(\nu, r)$ of $\alpha(\nu, r)$ can be recovered applying CBPF to $\widehat{\mathcal{Q}}$, with the same viewing conditions used in $\mathcal{I}$. That is, $\widehat{\alpha}(\nu, r)$ is the recovered e-CSF. Thus, the perceptual inverse quantizer or the recovered $\widehat{\alpha}(\nu, r)$ introduces perceptual criteria to Inverse Scalar Quantizer and is given by:

$$\widehat{\mathcal{I}} = \begin{cases} \sum_{s=1}^{n} \sum_{o=v,h,d} sign(\widehat{\omega_{s,o}}) \, \dfrac{\Delta_s \cdot (|\widehat{\omega_{s,o}}| + \delta)}{\widehat{\alpha}(\nu, r)} + (\widehat{c_n} + \delta) \cdot \Delta_n & |\widehat{\omega_{s,o}}| > 0 \\ 0, & \widehat{\omega_{s,o}} = 0 \end{cases} \tag{2}$$

For the sake of showing that the encoded VFWs are approximately equal to the decoded ones, that is $\alpha(\nu, r) \approx \widehat{\alpha}(\nu, r)$, we perform two experiments.

(a) JPEG2000 compressed at 0.59 bpp.



(b) JPEG2000-F-ρSQ compressed at 0.45 bpp.

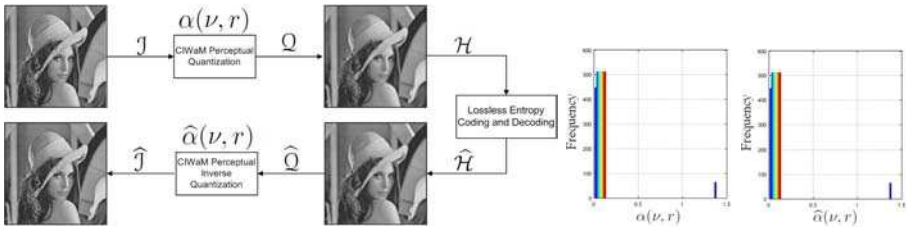
Figure 3. Examples of recovered images of Baboon.

Figure 4. The χSET image compression algorithm. Green blocks are the F-ρSQ and I-ρSQ procedures.

Experiment 1: Histogram of $\alpha(\nu, r)$ and $\widehat{\alpha}(\nu, r)$. The process of this short experiment is shown by Figure 5. Figure 5(a) depicts the process for obtaining losslessy both Encoded and Decoded visual weights for the 512×512 *Lena* image, channel Y at 10 meters. While Figures 5(b) and 5(c) shows the frequency histograms of $\alpha(\nu, r)$ and $\widehat{\alpha}(\nu, r)$, respectively. In both graphs, the horizontal axis represents the sort of VFW variations, whereas the vertical axis represents the number of repetitions in that particular VFW. The distribution in both histograms is similar and they have the same shape.

Experiment 2: Correlation analysis between $\alpha(\nu, r)$ and $\widehat{\alpha}(\nu, r)$. We employ the process shown in Fig. 5(a) for all the images of the CMU, CSIQ, and IVC Image Databases. In order to obtain $\widehat{\alpha}(\nu, r)$, we measure the lineal correlation between the original $\alpha(\nu, r)$ applied during the F-ρSQ process and the recovered $\widehat{\alpha}(\nu, r)$. Table 2 shows that there is a high similarity between the applied VFW and the recovered one, since their correlation is 0.9849, for gray-scale images, and 0.9840, for color images.

In this section, we only expose the results for the CMU image database.



(a) (b) (c)

Figure 5. (a) Graphical representation of a whole process of compression and decompression. Histograms of (b) $\alpha(\nu, r)$ and (c) $\widehat{\alpha}(\nu, r)$ visual frequency weights for the 512×512 image *Lenna*, channel Y at 10 meters.

Table 2. Correlation between $\alpha(\nu, r)$ and $\widehat{\alpha}(\nu, r)$ across CMU, CSIQ, and IVC Image Databases.

Image Database	8 bpp gray-scale	24 bpp color
CMU	0.9840	0.9857
CSIQ	0.9857	0.9851
IVC	0.9840	0.9840
Overall	**0.9849**	**0.9844**

Fig. 6 depicts the PSNR difference (dB) of each color image of the CMU database, that is, the gain in dB of image quality after applying $\widehat{\alpha}(\nu, r)$ at $d = 2000$ centimeters to the $\widehat{\mathcal{Q}}$ images. On average, this gain is about 15 dB. Visual examples of these results are shown by Fig. 7, where the right images are the original images, central images are perceptual quantized images after applying $\alpha(\nu, r)$ and left images are recovered images after applying $\widehat{\alpha}(\nu, r)$.

After applying $\widehat{\alpha}(\nu, r)$, a visual inspection of these sixteen recovered images show a perceptually lossless quality. We perform the same experiment experiment for gray-scale and color images with $d = 20$, 40, 60, 80, 100, 200, 400, 800, 1000 and 2000 centimeters, in addition to test their objective and subjective image quality by means of the PSNR and MSSIM metrics, respectively.

In Figs. 8 and 9, green functions denoted as F-ρSQ are the quality metrics of perceptual quantized images after applying $\alpha(\nu, r)$, while blue functions denoted as I-ρSQ are the quality metrics of recovered images after applying $\widehat{\alpha}(\nu, r)$. Thus, either for gray-scale or color images, both PSNR and MSSIM estimations of the quantized image $\mathcal{Q}$ decrease regarding d, the longer d the greater the image quality decline. When the image decoder recovers $\widehat{\mathcal{Q}}$ and it is perceptually inverse quantized, the quality barely varies and is close to perceptually lossless, no matter the distance.

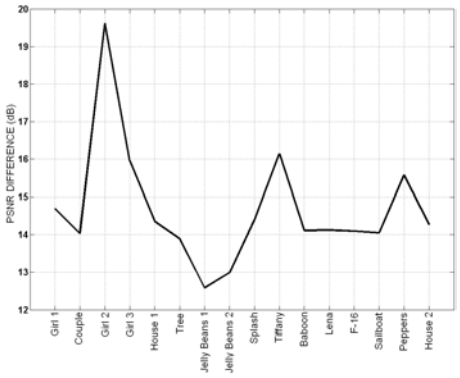


Figure 6. PSNR difference between $\widehat{\mathcal{Q}}$ image after applying $\alpha(\nu, r)$ and recovered $\widehat{\mathcal{I}}$ after applying $\widehat{\alpha}(\nu, r)$ for every color image of the CMU database.

5 Conclusions

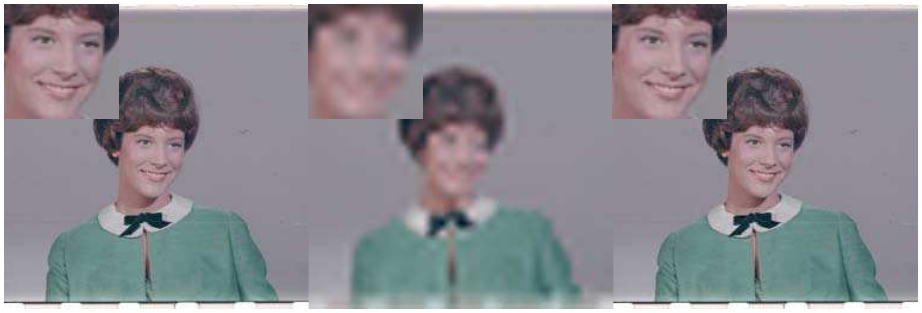
In this work, we defined both forward (F-ρSQ) and inverse (I-ρSQ) perceptual quantizer using CBPF. We incorporated it to Hi-SET, testing a perceptual image compression system χSET. In order to measure the effectiveness of the perceptual quantization, a performance analysis is done using thirteen assessments such as PSNR, MSSIM, VIF, WSNR or $\mathcal{NR}$PSNR, for instance, which measured the image quality between reconstructed and original images. The experimental results show that the solely usage of the Forward Perceptual Quantization improves the JPEG2000 compression and image perceptual quality. In addition, when both Forward and Inverse Quantization are applied into Hi-SET, it significatively improves the results regarding the JPEG2000 compression.

Acknowledgment

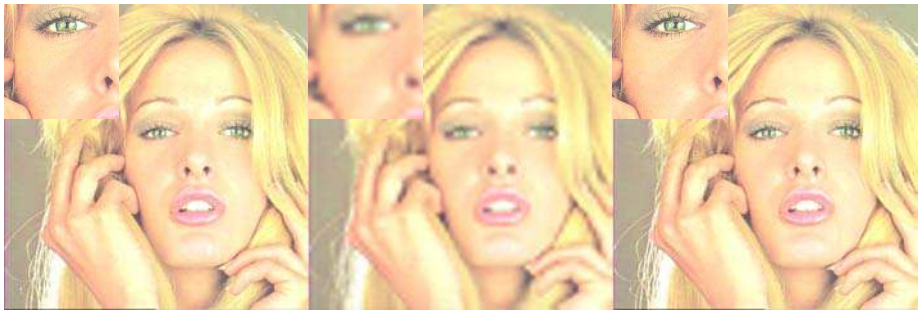
This work is supported by National Polytechnic Institute of Mexico by means of Project No. 20140096, the Academic Secretary and the Committee of Operation and Promotion of Academic Activities (COFAA), National Council of Science and Technology of Mexico by means of Project No. 204151/2013, and LABEX Σ-LIM France, Coimbra Group Scholarship Programme granted by University of Poitiers and Region of Poitou-Charentes, France.

References

1. M. Boliek, C. Christopoulos, and E. Majani, *Information Technology: JPEG2000 Image Coding System*, JPEG 2000 Part I final committee draft version 1.0 ed., ISO/IEC JTC1/SC29 WG1, JPEG 2000, April 2000.



(a) Girl 2

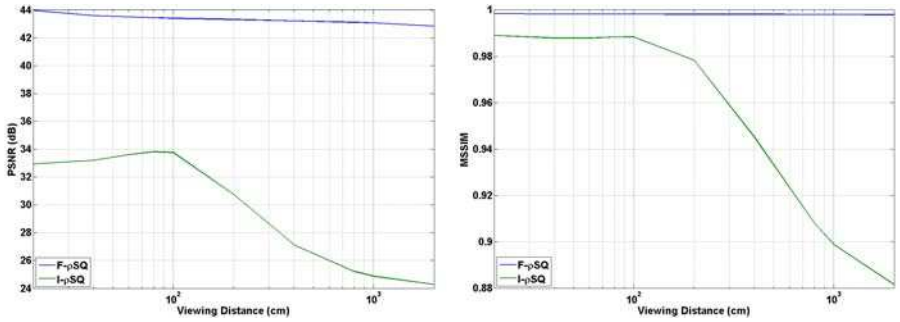


(b) Tiffany

(c) Peppers

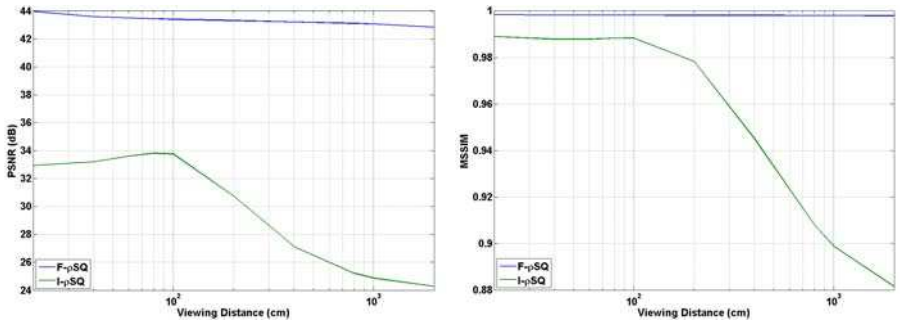
Figure 7. Visual examples of Perceptual Quantization. Left images are the original images, central images are forward perceptual quantized images (F-ρSQ) after applying $\alpha(\nu, r)$ at $d = 2000$ centimeters and right images are recovered I-ρSQ images after applying $\widehat{\alpha}(\nu, r)$.

2. S. I. P. I. of the University of Southern California. (1997) The USC-SIPI image database. Signal and Image Processing Institute of the University of Southern California. [Online]. Available: http://sipi.usc.edu/database/
3. C. Research, École Polytechnique Fédérale de Lausanne, and Ericsson. (2001) JJ2000 implementation in Java. Cannon Research, École Polytechnique Fédérale



(a) PSNR (b) MSSIM

Figure 8. PSNR and MSSIM assessments of compression of Gray-scale Images (Y Channel) of the CMU image database. Green functions denoted as F-ρSQ are the quality metrics of forward perceptual quantized images after applying $\alpha(\nu, r)$, while blue functions denoted as I-ρSQ are the quality metrics of recovered images after applying $\widehat{\alpha}(\nu, r)$.



(a) PSNR (b) MSSIM

Figure 9. PSNR and MSSIM assessments of compression of Color Images of the CMU image database. Green functions denoted as F-ρSQ are the quality metrics of forward perceptual quantized images after applying $\alpha(\nu, r)$, while blue functions denoted as I-ρSQ are the quality metrics of recovered images after applying $\widehat{\alpha}(\nu, r)$.

de Lausanne and Ericsson. [Online]. Available: http://jj2000.epfl.ch/

4. T. Mitsa and K. Varkur, "Evaluation of contrast sensitivity functions for formulation of quality measures incorporated in halftoning algorithms," *IEEE International Conference on Acustics, Speech and Signal Processing*, vol. 5, pp. 301–304, 1993.

Gaussian Noise Estimation Applied to Color Images

[1]F. Gamino, [1]A.J. Rosales, [1]F.J. Gallegos

[1]Av. IPN S/N, U.P.A.L.M., Co. Lindavista, C.P. 07738
f_gamino@hotmail.com; arosaless@ipn.mx; fgallegosf@ipn.mx

Abstract. An important field in Digital Image Processing is related to noise estimation, since it is essential for many image-processing procedures such as Filtering, Segmentation, Pattern recognition etc.; noise is inherent to any procedure and cannot be avoided since there are many types of noise from different sources, such as: image acquisition, image compression, transmission channel, etc. Therefore, we propose a novel Gaussian noise estimator using in first instance homogeneous areas and the whole image content by obtaining mean Euclidean distances between the noisy image and the image obtained by applying a mean filter to the noisy image. We validate the proposal against another algorithm, as well as their performance analysis for noise estimation and for a real application that consist in filtering noisy images using the CBM3D algorithm. Validations consist in using general-purpose images taken from the database BSDS500, which has low detailed images as well as high textured ones. Besides, our proposals is able to deliver better results for high noise levels.

Keywords: Gaussian Noise estimator, Color Images, Euclidean distance, Absolute and Angle distances.

1 Introduction

Basic procedures in the Gaussian noise estimation consist in estimating the standard deviation in order to determine the noise level present in the image, for our case in color images. For denoising image algorithms, knowing the type and the noise level to be treated is essential, which in most real applications is not common, thereby many methods of noise estimation have been proposed, Stanislav Pyatykh et. al. [1] consider three of them, which are:

1. Wavelet transformation
2. Image filtering
3. Preclassification of homogeneous areas

Another method proposed is found in [2] where Aihong Liuonly considers image filtering and preclassification of homogeneous areas.

Nevertheless, most of them have common criteria to take into account in their algorithms such as splitting noise, and image or modelling noise as a Gaussian distribution.

©L. Berthet, J. C. Chimal, E. A. Santos, E. Castillo
Advances in Computer Science
Research in computing Science 81, 2014 pp. 35-46

The present paper exposes two noise-estimation algorithms for Additive White Gaussian Noise (AWGN) [3,4] modeled by the equation (1), where Ar is the noisy image, A is the original image and $n(\mu, \sigma^2)$ is the AWGN, with mean $\mu = 0$ and variance σ^2. The proposed algorithms in this paper are based in filtering and splitting noise in the image, the first proposal uses the whole image and the second uses preclassification of homogeneous areas. Since these algorithms are improvements of an algorithm presented by Rosales-Silva et.al. in [5], a quickly review is needed, this explanation is given in section 2.

$$Ar = A + n(\mu, \sigma^2). \tag{1}$$

To validate the proposed noise estimators, we implemented the algorithm designed by Xinhao Liu et.al. [6] ("the reference algorithm") which uses a technic based on patches selection, and a "tuning" process according to the image in which the estimator is been applied, however this algorithm process R,G,B channels individually, so it obtains three different results for each estimation, so it does not take into consideration the correlation in the R,G,B color channels.
Xinhao Liu et.al. in [6] states, "even if the true noise level is estimated ideally, the denoising algorithm might not achieve the best performance.".

In order to make comparisons to our proposals, the algorithm Color Block Matching 3D "CBM3D" presented in [9] by Dabov et.al. is implemented. The algorithm CBM3D consists of selecting sets of similar blocks of pixels which collaborate in the denoising procedure, this means that, each block provide information for the robustness in the denoising algorithm.

The paper is organized as follows: in Section 2 is introduced the theoretical background in which our algorithms are based; Section 3 exposes the detailed methodology of the two algorithms proposed; Section 4 exposes analysis and comparisons between the proposed and reference algorithms using images from the Berkeley Segmentation database "BSDS500" [8]; Section 5 presents the conclusions.

2 Distance Estimation Criteria

Our proposed algorithms improve the results of a previous work treated in [5], the previous work algorithm "basic algorithm" consists in three steps:

1. Filter the noisy image (Ar) by a mean filter [10] using a window processing of 3X3 pixels, so a filtered image (Af) is obtained.

$$Af = MeanFilter(Ar). \tag{2}$$

2. Obtain the angles $(Ang_{i,j})$, Euclidean $(D_{eu_{i,j}})$and absolute distances $(D_{ab_{i,j}})$, between pixels in the same spatial position $(i = 0,1,2, \ldots., M;\ j = 0,1,2, \ldots., N;$ between the filtered image and the noisy image, assuming them to be three dimensional (R, G, B) points.

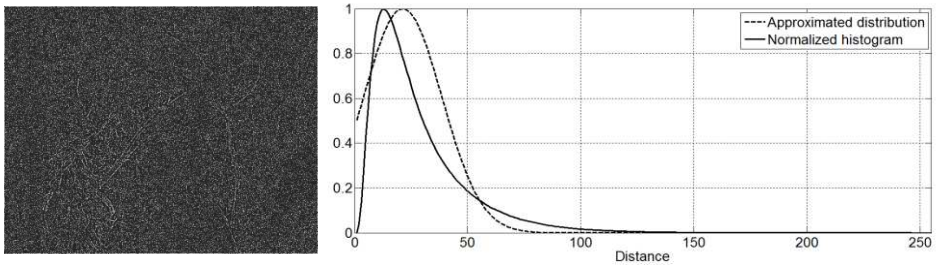
$$D_{ab_{i,j}} = \left|\left(Ar_{R_{i,j}} - Af_{R_{i,j}}\right) + \left(Ar_{G_{i,j}} - Af_{G_{i,j}}\right) + Ar_{B_{i,j}} + Af_{B_{i,j}}\right|, \quad (3)$$

$$D_{eu_{i,j}} = \sqrt{\left(Ar_{R_{i,j}} - Af_{R_{i,j}}\right)^2 + \left(Ar_{G_{i,j}} - Af_{G_{i,j}}\right)^2 + (Ar_{B_{i,j}} - Af_{B_{i,j}})^2}, \quad (4)$$

$$Ang_{i,j} = cos^{-1}\left(\frac{Ar_{R_{i,j}}.Af_{R_{i,j}} + Ar_{G_{i,j}}.Af_{G_{i,j}} + Ar_{B_{i,j}}.Af_{B_{i,j}}}{\sqrt{Ar_{R_{i,j}}^2 + Ar_{G_{i,j}}^2 + Ar_{B_{i,j}}^2}.\sqrt{Af_{R_{i,j}}^2 + Af_{G_{i,j}}^2 + Af_{B_{i,j}}^2}}\right). \quad (5)$$

3. Compute the mean and the standard deviation for angles, Euclidean and absolute distances obtained from Eq.(3), Eq.(4), Eq.(5) respectively assuming that a Gaussian distribution describes the data to be taken into account.

As an example, the results obtained from the proposed algorithm applied in the well-known image Lena by contaminating it with a noise level of 25, are shown in fig 1. Even when Absolute and angle distances were analyzed, for commodity only the process for the Euclidean distances (fig 1a) is shown, fig 1a was obtained by applying eq. (4) to each pixel, between noisy image and filtered image. Euclidean distances histogram normalized and its approximated Gaussian distribution in fig 1b are obtained in order to calculate the mean Euclidean distance and its standard deviation, this parameter is supposed to be equal to the noise level introduced in the original image in an artificial way using Matlab functions.



(a) (b)

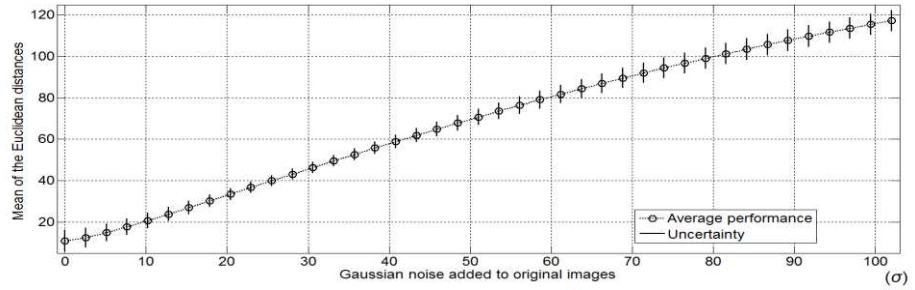
Fig. 1. a) Euclidean distances between pixels in noisy and filtered images, b) normalized histogram and approximated Gaussian distribution for Euclidean distances in (a).

The previous method has low accuracy because it assumes that filtered images have no noise, which is not true, for most of the real images to be processed. The Gaussian noise added artificially has an standard deviation of 25 and after applying the basic algorithm to estimate the noise level was obtained to be equal to 17.31 by using the histogram distances. This result shows that the estimated value is different from the original noise level artificially added; this is the main objective treated in this paper.

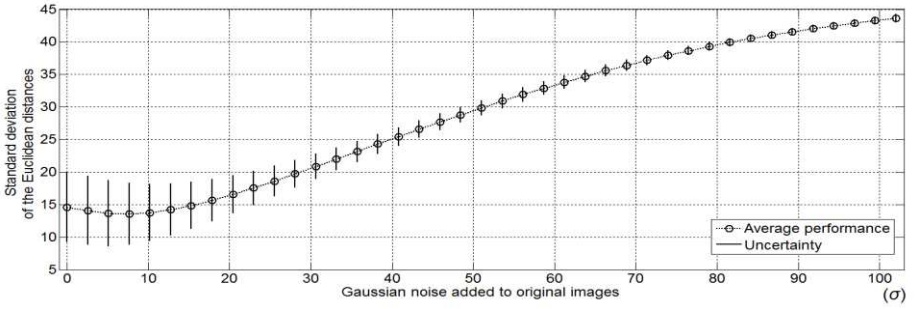
2.1 Basic algorithm performance

After several experiments, where different noise levels were added to images from BSDS500, it was found a relationship between the results obtained from the basic algorithm presented in section 2 and the standard deviation of the AWGN previously added to the images. This is shown in the fig. 2, only the analysis for mean Euclidean distances is shown because after some analysis, it was demonstrated to have a better performance for Gaussian noise estimation, also in fig 2a could be seen that the relation between the noise level added to the image and the mean of the Euclidian distances tends to be a monotonic function which could be modeled as an straight line, while in fig 2b its possible to see that the standard deviation of the Euclidian distances result in a non-monotonic function which should be modeled as a polynomial function.

In fig 2a and 2b, is shown the average performance and the uncertainty for each point from applying the basic algorithm to all images in BSDS500, adding different amounts of noise from 0 to 100 in increments of 2 in its standard deviation.



(a)



(b)

Fig. 2. Average performance for all images in BSDS500 "doted and circles line" including uncertainties "black continuous lines" for each point analyzed, a) using the mean Euclidean distance, b) using the standard deviation of the Euclidean distances.

So, our proposed algorithms were based in the results presented in fig 2 a, this proposals are explained in section three, on the basis of the basic algorithm presented at the beginning of this section.

3 Proposed Algorithms for Noise Estimation

As it could be seen in fig 2a, the average behavior can be modeled as a straight line just by applying a linearization method, while in the fig 2b this does not apply because in that case the average performance is a non-monotonic function and it produces that for certain amounts of noise two results could be obtained from the algorithm where only one of them is the correct output.

3.1 First Proposal (EstGl)

Let's suppose that the red line in the fig 2a is an straight line σ represented by eq. (6), which is equal to mean Euclidean distance "v" plus-minus its uncertainty "$\Delta\sigma$", in order to obtain a value representing the noise level from this equation its necessary to approximate it to the ideal case, which is when the value obtained is exactly the same noise level which was added artificially to the processing image. So this is an straight line with slope equal to one, first of all it is necessary to adjust σ to cero by taking from it, the minimum value "v_{min}"; then normalize it by dividing for "v_{max}" and finally multiply the value resulted by the maximum value considered for the estimator "x_{max}" which we proposed as 100 in order to avoid arithmetic overflow (arithmetic overflow will be discussed in Section 5). This procedure is simplified as "$F = (x_{max}/(v_{max} - v_{min}))$". After testing this adjustment for the basic algorithm, F was experimentally obtained equal to 0.95 and v_{min} equal to 10.96.

$$\sigma = v \pm \Delta\sigma, \tag{6}$$

$$\sigma_{aprox} = Fv - Fv_{min} \pm F\Delta\sigma. \tag{7}$$

σ_{aprox} represents a line whit the same performance shown in fig 2a, so in order to have a better estimator's response, it is proposed a piecewise linearization, developed by applying a least squares linearization, for three segments. Intervals for every segment's limits which resulted in lower errors were proposed. Then making the assumption that eq. (7) represents an straight line equal to $y_n = m_n x + b_n$; where x represents the AWGN added to the image, m_n is the slope in the *n*-th segment and $b_n$is the point where the line y_n crosses *y* axis, the term $\Delta\sigma$ could be omitted in order to avoid uncertainties propagation. So after Isolating *x* in algebraic treatment, we obtain the equation (9), and to simplify we propose the use of B_n and K_n variables.

$$(v - v_{min})F = m_n x + b_n. \tag{8}$$

$$x = \frac{vF}{m_n} - \frac{Fv_{min}}{m_n} - \frac{b_n}{m_n}, \tag{9}$$

$$K_n = \frac{F}{m_n}\ , \tag{10}$$

$$B_n = -\frac{Fv_{min}}{m_n} - \frac{b_n}{m_n}. \tag{11}$$

The term $\Delta\sigma_{est}$ in eq. (12) is the uncertainty of the estimator which is calculated under several experiments.

$$\sigma_{est} = K_n v + B_n \pm \Delta\sigma_{est}. \quad (12)$$

Table 1. Values for constants K_n, B_n and intervals for each segment.

n-th Segment	Proportionality constant K_n	Adjust to cero constant B_n	Values obtained from eq. (7)
1	1.02	-32.40	Lower than 13
2	0.82	-7.21	From 13 to 67
3	1.13	-10.41	Higher than 67

The estimator proposed will be called EstGl, given by eq. (12), although a specific process is necessary to determine the value of the constants to be used in eq. (12), this process is explained in the next blocks diagram.

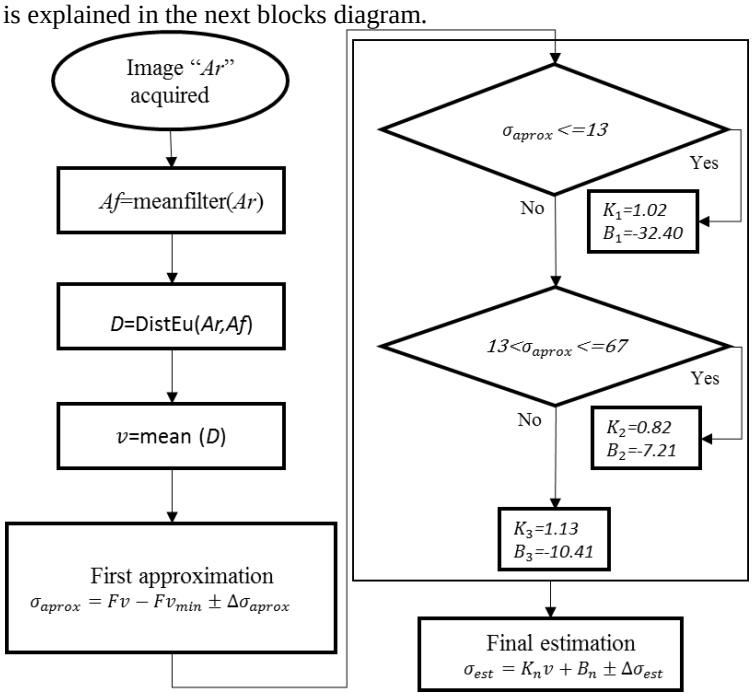


Fig. 3. Proposed Estimator EstGl.

3.2 Second Proposal (EstBl)

An adjustment on EstGl called EstBl consists in an homogeneous area pre-selection, it consists in dividing the image in blocks of 13X13 pixels and then apply the EstGl for each block as if they were complete images, taking as the result obtained by the estimator the smallest value obtained from the blocks which represents the most homogeneous area in the image, where, as mentioned in [6] is more likely to obtain the real

noise level, this is mainly because homogeneous areas do not contain details which could be mistaken for noise in the noise estimation algorithm.

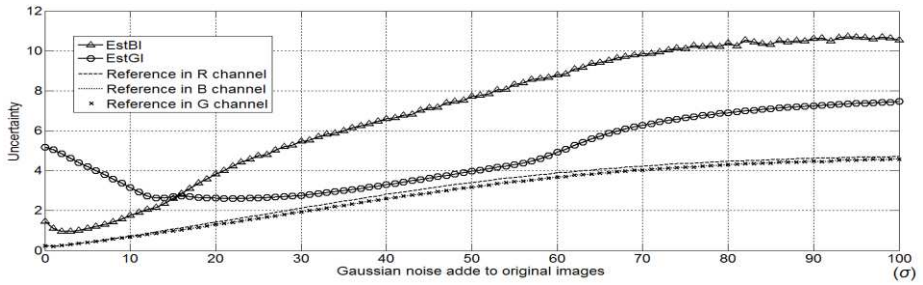
4 Results

In order to probe the performance of the proposed algorithms two tests were applied, the first one is for noise estimation (Section 4.1), which consist in contaminating the images from BSDS500 by a known noise level, then estimating the noise level using the proposed algorithms and "the reference algorithm", the second one in Section 4.2 consists in contaminating the images from BSDS500 by a known noise level and filter those noisy images using the CBM3D algorithm taking into account the estimations obtained by the proposed algorithms, "the reference algorithm" and the real noise level.

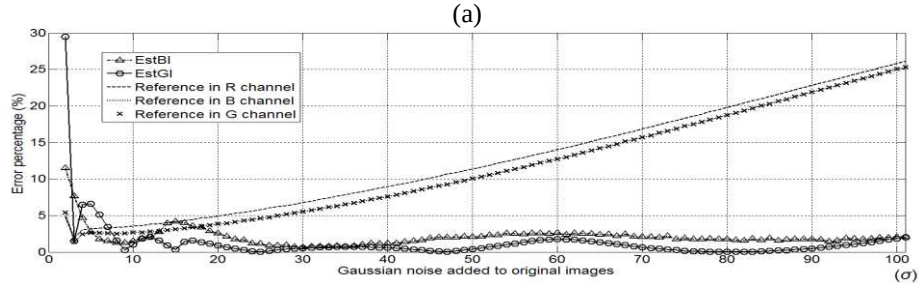
4.1 Performance Results in Noise estimation

We use all images in BSDS500, and the implementation of a comparison to "the reference algorithm", the proposed methodology evaluation consists on the next steps:

1. Acquire an original image from the database BSDS500.
2. Add an amount of AWGN to the original image.
3. Estimate the noise level from the noisy image.



(a)



(b)

Fig. 4. Display of uncertainty (a) and error percentage (b), for proposed estimators and reference estimator. EstGl "continuous line with circles", EstBl "continuous line with triangles", estimation for red by reference estimator "discontinuous line", estimation for green by reference estimator "x-points" and estimation for blue by reference estimator "dotted line".

This procedure was applied for all images in BSDS500 and for a range of amounts of noise "from 0 to 100 in its σ in one by one increments", after obtaining results for proposed estimators and reference, their average performance was calculated as well as their uncertainty and mean percentage error which are shown in fig 4.

In fig 4 is shown the uncertainty (a) and error percentage (b), in the proposed algorithms we have a lower error for amounts of noise greater than 10, the proposed algorithms have higher accuracy for amounts of noise from 10 to 100 but with lower precision for amounts of noise from 0 to 100, this will be discussed in Section 5.

So, after proving that our proposed algorithms have an acceptable performance for noise estimation a second test was proposed which consists in filtering images artificially contaminated, this test and the results obtained from it are explained in Section 4.2.

4.2 Filtering images

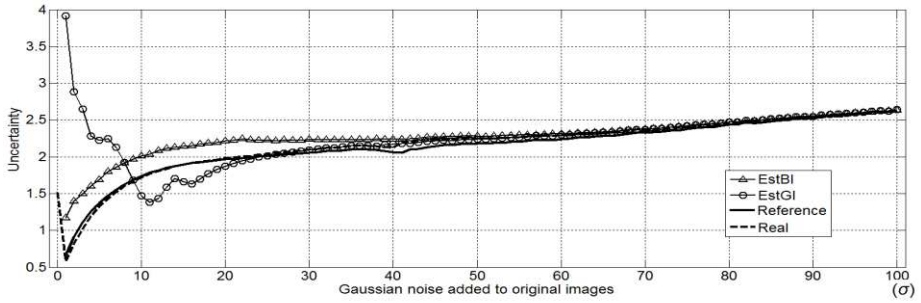
For some filtering algorithms is necessary to know a priori the amount of AWGN contained in images, one of them is the CBM3D presented in [9] and which also has been used as reference because it has proved to be one of the best denoising algorithms. Therefore, another comparison between proposed estimators and "the reference estimator" performance, is made by adding a fourth step in the procedure presented in section 4.1, which is:

4. Filter the noisy images by CBM3D, taking into account the real and estimated amounts of noise.

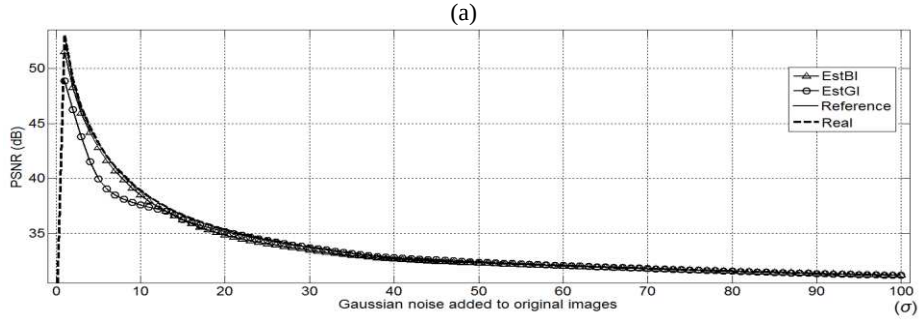
As mentioned in Section 4.1, "the reference algorithm" estimates amounts of noise for each channel in the R,G,B color space model, so in order to have just one estimation from the reference algorithm, a mean value was obtained from these results calculating the average of the noise values computed for every channel.

Results for filtered images were evaluated using the Peak Signal-to-Noise Ratio (PSNR) eq. (13) where I and k represents the evaluated and reference images respectively; the average PSNR for different noise levels and its uncertainty considering all results obtained are displayed in fig 5.

$$PSNR = 10\,log_{10}\left(\frac{255^2}{\frac{1}{3xMxN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\sum_{\beta=R,G,B}\left\|I_\beta(i,j)-k_\beta(i,j)\right\|^2}\right). \quad (13)$$



(a)



(b)

Fig. 5. Display of uncertainty (a) and PSNR (b), for filtering using CBM3D considering results obtained from EstGl "continuous line with circles", EstBl "continuous line with triangles", "the reference algorithm" "continuous line" and real "discontinuous line".

In fig 5 are displayed results obtained from the analysis of 439 images, even when BSDS500 has 500 images only 439 images where used for this analysis, since "the reference algorithm" algorithm deliver complex numbers as estimation after processing 61 images at low noise levels lower than 5.

5 Discussion

In section 4.1 uncertainty and error percentage are displayed in fig 4 for results obtained from proposed and "the reference algorithm"; as mentioned, proposed algorithms have lower error with greater uncertainty especially EstBl.

However uncertainty increases as the noise level increase this effect could be a consequence of arithmetic overflow error which "cuts" and deforms distribution's tails, for example if the AWGN added has a big standard deviation, representing the noise level and we assume that after this process is possible to split the image and AWGN, both of them would have been modified because of the arithmetic overflow error which in this case leads to the consequence of the impossibility of representing some values obtained by adding noise to an image.

In section 4.2 uncertainty and PSNR criteria are displayed in fig 5, it is remarkable that even when proposed algorithms "EstGl and EstBl" have lower errors in comparison with "the reference algorithm" (fig 4b) , they have a similar average performance in the quality of the filtered images measured by the PSNR (fig 5b), considering estimations by all analyzed methods and the "real noise level", except for EstGl when the noise level is lower than 10.

Also it is important to say that uncertainty shown in fig 5a, is similar for estimations made by EstBl, and "the reference algorithm" and are also similar to the results obtained by considering the real noise level.

6 Conclusion.

Two algorithms were proposed the EstBl and the EstGl, was demonstrated that they have an acceptable performance, in noise estimation for images from the BSDS500 and in real applications.

However, arithmetic overflow errors are still a problem mainly for high noise levels (for our consideration greater than 35) as shown in the results of the fig 4 and the fig 5.

Another important fact is that these algorithms were proposed and tested for general-purpose images. So it is possible to apply the same principle for an specific type of images in order to have a better estimation, the variation would be the requirement of having to calculate all the constants needed, F, $vmin$, K, B and proposing limit intervals in order to obtain better results for noise level estimation as dictated in Table 1.

Acknowledgments

The authors thank to the National Polytechnic Institute of Mexico (Instituto Politécnico Nacional de México), and the National Council of Science and Technology of Mexico (CONACYT) for their support to make possible this work.

References

1. S. Pyatykh, J. Hesser, L. Zheng, Image Noise Level Estimation by Principal Component Analysis, IEEE Transactions on Image Processing, vol.22, no.2, pp.687,699, Feb (2013).
2. Aihong Liu, A Fast Method of Estimating Gaussian Noise, Information Science and Engineering (ICISE), 2009 1st International Conference on, pp.441,444, 26-28 Dec. (2009).
3. Russo, F, A method for estimation and filtering of Gaussian noise in images, Instrumentation and Measurement Technology Conference, 2002. IMTC/2002. Proceedings of the 19th IEEE , vol.2, no., pp.1689,1694 vol.2, 2002.

4. Wei Liu, Weisi Lin, Additive White Gaussian Noise Level Estimation in SVD Domain for Images, Image Processing, IEEE Transactions on , vol.22, no.3, pp.872,883, March 2013.
5. A. J. Rosales-Silva, F. J. Gallegos-Funes, I. Bazán-Trujillo, and A. Ramírez García, Robust fuzzy scheme for Gaussian denoising of 3D color video, 2014,EURASIP Journal on Image and Video Processing,20,2014.
6. Xinhao Liu, M. Tanaka, M. Okutomi, Single-Image Noise Level Estimation for Blind Denoising, Image Processing, IEEE Transactions on, vol.22, no.12, pp.5226,5237, Dec. 2013.
7. Dong-Hyuk Shin, Rae-Hong Park, Seungjoon Yang, Jae-Han Jung, Block-based noise estimation using adaptive Gaussian filtering, Consumer Electronics, IEEE Transactions on, vol.51, no.1, pp.218,226, Feb. 2005.
8. D. Martin. C. Fowlkes, D. Tal and J. Malik, A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics, pp 416—423, July 2001.
9. K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering, Image Processing, IEEE Transactions on , vol.16, no.8, pp.2080,2095, Aug. 2007.
10. I. Pitas and A. N. Venetsanopoulos, Nonlinear Digital Filters: Principles and Applications. Boston, MA: Kluwer Academic, 1990.

Extended Firefly Algorithm and its Application in Unsupervised Classification of Mixed and Incomplete Data

Yenny Villuendas-Rey,[*1], José Felix Cabrera Venegas[1], and Jarvin Alberto Anton Vargas[1]

[1]University of Ciego de Ávila, Road to Morón km 9 ½, Ciego de Ávila, Cuba
{yenny, josefelix, janton@cav.uci.cu}

Abstract. Bioinspired Algorithms are ones of the most powerful metaheuristics used to solve optimization problems because its low computational cost and ability to explore a wide space of solutions in a little time. Unsupervised Classification may be comprehended as an optimization problem where the goal is to find the right data clustering among the numerous ways to do data clustering. In this paper we propose a new extension to Firefly Algorithm FA. The propose method is based in modifications to the original metaheuristic and the redefinition of artificial fireflies and the objective function for adjust the algorithm to solve a general problem. This approach was applied to Unsupervised Clasification with mixed and incomplete data. Experimental analysis with other algorithms using repository databases shows that our approach is able to find compact and separate clusters as well as to estimate the natural structuration of data.

Keywords. unsupervised classification, bioinspired metaheuristics, firefly algorithm (FA).

1 Introduction

Unsupervised Classification consists on obtaining a partition of the dataset in such way that the objects belonging to the same group be more similar to each other that with regard to the objects of other groups. In a general way, the algorithms of Unsupervised Classification are based on some approach that reflects how good it is a certain partition of the data [1].

Several authors have considered Unsupervised Classification as a problem of optimization, where the goal is to obtain the partition that maximizes or minimized the objective desired (the quality of the obtained structuring). Therefore, it has been carried out some proposals for the application of algorithms of optimization to the solution of this problem in particular. Contrary to most of the problems of optimization, where it is necessary to find an n-dimensional vector that satisfies the objective to optimize, in the case of Unsupervised Classification we have the space of solutions formed by all the possible partitions to obtain. In a similar way, it happens in other

©L. Berthet, J. C. Chimal, E. A. Santos, E. Castillo
Advances in Computer Science
Research in computing Science 81, 2014 pp. 47-57

domains of the Artificial Intelligence, such as the prototypes selection, rules generation, etc. [2]. Therefore, it is necessary to outline the problem of optimization in a general way, without assuming vectorial spaces.

Recently, bioinspired algorithms based in swarm intelligence have been applied successfully to the problem of the Unsupervised Classification [3], [4], [5], although only for numeric data, not being proposals for Mixed and Incomplete Data (MID). In this work, we will propose an extension to the Firefly Algorithm (FA) for solve general optimization problems, and we will apply it to the Unsupervised Classification of MID.

2 Extended Firefly Algorithm to obtain Groups in Mixed and Incomplete Data

Bioinspired Metaheuristics have proved their efficiency in Unsupervised Classification [6], [7], [8]. However, most algorithms only work only with numerical data. In this section we explain the metaheuristic Firefly Algorithm (FA) proposed by Xin-She Yang [9] for numerical optimization, and its extension to solve the problem of Unsupervised Classification where objects with mixed and incomplete data descriptions have to be arranged in groups.

2.1 New Metaheuristic based in an Extension to Firefly Algorithm

The social behavior of the fireflies has focused the attention of many computer scientists, basically regarding to the light they flash. The flashing light can be formulated in such a way that it is associated with the objective function to be optimized, which makes it possible to formulate new optimization algorithms. One of these techniques is a Firefly Algorithm (FA) for multimodal optimization applications developed by Xin-She Yang in 2009 [9]. FA was proposed originally for numerical optimization, it can be modified for obtaining of groups of objects in Unsupervised Classification of MID. These modifications are addressed below.

It is known that the selection of appropriate values for the parameters of the algorithms is crucial for a good performance of them. In this case, the algorithm FA consists of four fundamental parameters: β, γ (attractiveness and variation of the attractiveness) and the quantity of fireflies η. Investigations carried out by Lukasik and Zak [10] on the FA allows concluding that the best values for these parameters are: β =1, γ =1 and η varying between 15 and 50.

These conclusions allow simplifying the movement of the firefly defined in [10], as follows:

$$X_i = x_i + \beta(x_j - x_i + \alpha\varepsilon_i), \text{ with } \beta = 1$$

$$X_i = x_i + x_j - x_i + \alpha\varepsilon_i$$

$$X_i = x_j + \alpha\varepsilon_i \quad (1)$$

Therefore, the movement of firefly i transforms into a modification of the position of firefly j. In this idea we base ourselves to develop the extension of the algorithm. In a general way, a modification of the position of the brighter firefly can be considered like a "perturbation" of this firefly. With the objective of extending the original algorithm, we consider that each firefly consists, more than in a certain position, in a solution candidate to the problem of optimization we want to solve.

Furthermore, the characteristics of the fireflies in the original algorithm [10] are redefined in the following way:

1. Each firefly represents a candidate valid solution to the problem of optimization we want to solve.
2. All the fireflies are unisex, that means that a firefly can attract other fireflies without considerate its sex.
3. The attractiveness of the firefly is proportional to its brightness; in this case the less bright ones will be the result of a perturbation of brighter firefly, being the last one the more attractive. If there is not a firefly brighter that the current firefly, then the current firefly is perturbed because it is not attracted by any firefly.
4. The brightness of a firefly is affected or it is determined by the objective function that will be optimized.

Also, we assume that exists a way of "perturbing" this solution. Then, the "movement" of fireflies will be given for (2), instead of equation (1):

$$X_i = Perturb\ (x_j) \tag{2}$$

Each metaheuristic must implement a balance between search in the solutions space (exploration) and intensification of the good solutions (exploitation). In the particular case of the algorithms based on swarms of fireflies, the intensification of the good solutions is given by the realization of perturbations in the best fireflies. This procedure reduce the exploration of the search space by the perturbation of the best firefly we have found, reducing the possibilities to explore areas of the search space that are far from the area represented by the best fireflies.

To solve this problem, we include a new characteristic to the fireflies: their Time of Life. The Time of Life of firefly is a parameter of the algorithm, and it defines the number of iterations each firefly will "live". The age of the fireflies is initialized in 1 when the fireflies are generated initially, and it is increased each iteration. Every time that a firefly is perturbed, her age is again 1. Then, if a firefly is not perturbed during a number of iterations defined (it exceeds its time of life), we consider that this firefly dies, and it is replaced for other firefly generated according to the generation procedure used by the algorithm

This new characteristic is able to take out the algorithm of local optimal solutions and, at the same time, it improves the possibilities to explore big areas to the search space. Although, the modifications introduced in the metaheuristic allow a bigger exploration of the search space, it is possible to lose good solutions, even optimal solutions, due to the "death" of the fireflies. In Fig. 1, we show the pseudo code of the Firefly Algorithm with the modifications discussed in this section.

Extended Firefly Algorithm

Input: η: quantity of fireflies
A: fireflies attractiveness (objective function)
I: quantity of iterations
t: time of life of fireflies (age)

Output: *Best_F*: Best firefly (solution to the optimization problem)

Stage 1. Fireflies Inicialization (creating candidate solutions)

for $i = 1$ **to** η
  F_i = *Generate_Firefly()*
  $F_i.age = 1$
end for *i*
Hall_Fame = $arg\ \max_{i=1\ldots\eta}\{A(F_i)\}$

Stage 2. Shine of Fireflies (exploration the search space)

Iteration =1;
while (*Iteration* < *I*)
  for $i = 1$ **to** η
    for $j = 1$ **to** η
      if ($A(F_j) > A(F_i)$)
        F_i = *Perturb*(F_j)
        $F_i.age = 1$ **end if**
    end for *j*
    if ($F_i.age > t$)
      Consider_to_Fame(F_i)
      Replace_Firefly(F_i)
    else
      $F_i.age$++
    end if-else
  end for *i*
  Best_F = $arg\ max_{i=1\ldots\eta}\{A(F_i)\}$
  Best_F = *Perturb(Best_F)*
  Iteration++
end while
if (*A(Hall_Fame)* > *A*(*Best_F*))
  Best_F = *Hall_Fame* **end if**

Fig. 1. Pseudo code of extended firefly algorithm (EFA).

Another important feature is the fireflies die. For example, if we have a good solution in the 2nd iteration of the algorithm, and the time of life of the firefly is 3 iterations, this optimal solution won't be improved, and possibly, we will get lost the solution due to the firefly die. To solve this problem, we consider introducing a Hall of Fame in the metaheuristic. The Hall of Fame will be dedicated to store the information of the best firefly (solution of the problem) founded.

When a firefly dies, it will be considered to integrate the Hall of Fame (*Consider_to_Fame*). This means that the firefly in question (candidate) will be compared with the existent firefly in the Hall, and if it is better than the last one, the existent firefly in the Hall of the Fame will be replaced by the candidate firefly.

The new metaheuristc propose above, allow finding the solution of problems in a general way, not only restricted to numeric search spaces. Also, the introduction of Time of life and the Hall of Fame can improve the quality of the obtained solution. In the following section is analyzed the application of this metaheuristic in Unsupervised Classification of Mixed and Incomplete Data (MID).

2.2 Extended Firefly Algorithm applied to Unsupervised Classification of MID (EFAC)

The process of Unsupervised Classification can be seen as a combinatorial problem of optimization. In this case, the space of solutions is given by all the possible ways to create groups of objects, and the function to optimize is the quality of the obtained groups. We will call to the approach we will explain next: Extended Firefly Algorithm for Clustering (EFAC). With the goal of applying the proposed metaheuristic in Unsupervised Classification of MID, it is necessary firstly to define how the fireflies will be modeled and which attractiveness will be used (objective function). Also, we need to define how the firefly perturbation will be made.

First, we have to delimit the problem to the domain of Restrict Algorithms of Unsupervised Classification. This means that we need to know the value of k, the number of groups we want to obtain.

For solving our problem, a firefly will be a way to group the data objects (a candidate clustering). To model this clustering, each firefly represents the k centers of the k groups and is built in the following way:

$$F_i = (\bar{c}_1, \bar{c}_j \ \ldots, \bar{c}_k) \tag{3}$$

where represents the center of the group *j*, in the firefly *i* (*i* candidate clustering).

To allow the handling of MID, instead of using the average of the objects of the group like center, we select as center of a group the object that minimizes the disimilarity with rest of objects of its group:

$$\bar{c}_j = arg\, min_{x,y \epsilon C_j}\{d(x,y)\} \tag{4}$$

To compute the distance between two objects $d(x,y)$ with MID attributes, we used the HEOM dissimilarity (equation 5) proposed by Wilson and Martínez [11].

$$HEOM(x,y) = \sqrt{\sum_{a=1}^{m} d_a(x_a, y_a)}$$

$$d_a = \begin{cases} 1 & unknown\ attribute \\ overlap(x_a, y_a) & nominal\ attribute \\ diff(x_a, y_a) & otherwise \end{cases}$$

$$overlap(x_a, y_a) = \begin{cases} 0 & if\ x_a = y_a \\ 1 & elsewhere \end{cases}$$

$$diff(x_a, y_a) = |x_a - y_a| / (max_a - min_a) \quad (5)$$

max_a and min_a are the maximum and minimum values of attribute a, respectively.

Other important element in the algorithm is the attractiveness or objective function. This function measures the fitness of each clustering (solution) and permit to know which clustering is the best. The objective function we use in our investigation is the Dunn index (equation 6), defined in [12]. The Dunn index for a clustering is the quotient between the smallest distance among two groups, and the size of the biggest group. A high value of Dunn index for a clustering means we have more compact and more separate groups [13].

$$D = \frac{\min\limits_{i=1..k,\, j=1..k,\, i \neq j} \{d(c_i, c_j)\}}{\max\limits_{i=1..k} \{\Delta(c_i)\}} \quad (6)$$

Another element of the algorithm is the generation of fireflies. In our case, we use a random generation process. We select of the database of objects *k* elements that will be the centers of the groups. This strategy of random selection of the centers of the groups allows exploring a wide area of the search space, and it allows maintaining the diversity in the group of fireflies.

Finally, we defined how to perturb the fireflies. For this, we develop a method to produce a new solutions form a current solution inspires by the strategy of mutation of the Genetic Algorithm proposed for [7]. Our perturbation consists on replacing randomly one of the centers of the groups for another object in a random way.

The selection of the center that will be replaced is realized randomly. For this, a random number is generated indicating what center will be changed. Then, another random number is generated representing the object of the dataset that will replace the selected center. This strategy allows the exploration of wide areas of the search space, and it avoids a premature convergence of the algorithm.

The Fig. 2 shows an example of a firefly perturbation. In the figure, we represent the original firefly and below, the perturbed firefly. In de perturbed, the center of the second group is replaced randomly by another object of the group.

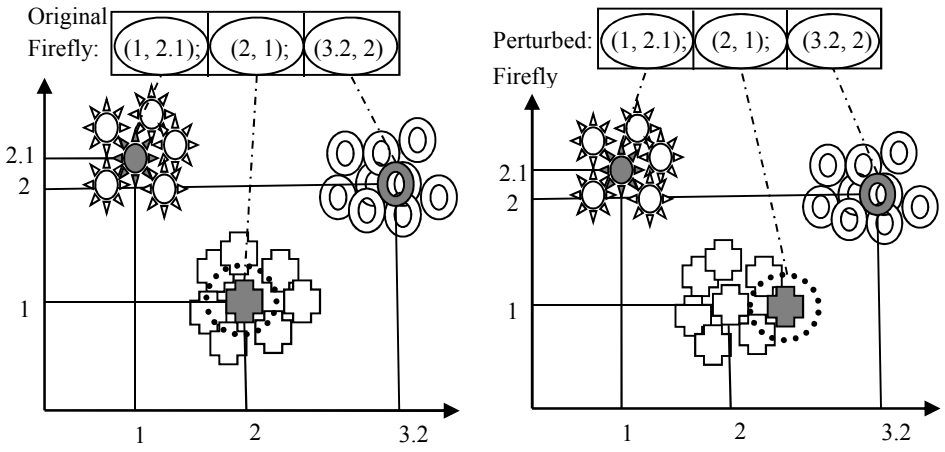


Fig. 2. Representation of a firefly perturbation.

3 Experimentation and Results

To check the performance of the proposed algorithm (EFAC) we developed experiments with six algorithms used in Unsupervised Classification of MID. These algorithms were divided in two groups. In the first we have the algorithms based in kmeans: KP [14], KMSF [15], AD2011 [16]. In de second group we have algorithms that use hierarchical and metaheuristic approaches: HIMIC [17], CEBMDC [18] and AGKA [7]. In the comparison were used seven data sets of Mixed and Incomplete Data (MID) from the UCI Repository [19], in Table 1 are shown the description of each data set.

Table 1. Data sets description.

Data sets	Categorical Features	Numerical Features	Classes
colic	15	7	2
dermatology	1	33	6
heart-c	7	6	5
hepatitis	13	6	2
labor	6	8	2
lymph	15	3	4
tae	2	3	3

To measure the obtained results we used the classes of the data sets as the real clustering. For each obtained clustering we calculated its Entropy [20]. Entropy is an external validity index defined by:

$$E = -\sum_i \frac{|C_i|}{|O|} \sum_j \frac{\left|\left\{o \in C_i \mid \alpha(o) = l_j\right\}\right|}{|C_i|} \log\left(\frac{\left|\left\{o \in C_i \mid \alpha(o) = l_j\right\}\right|}{|C_i|}\right) \quad (9)$$

Where O is the set of objects (data set), C is the obtained clustering, C_i is the group i, l_j is the class j and α (o) it is the class of the object o.

Entropy gives a measure of the difference between the groups obtained by an unsupervised classifier and the original classes in a data set. According to the above-mentioned Entropy measures the grade of disorganization of the original groups (classes in the data set) with regard to obtained groups (clustering results). Therefore, when the value of Entropy is smaller, the groups will be more similar to the data set classes.

We applied the different algorithms for each data set. Then we computed the Entropy index of each obtained clustering. In each algorithm, the quantity of groups was established as the quantity of classes for each data set. The dissimilarity used in the experiments for all the algorithms was HEOM [11] (equation 8). This dissimilarity has been applied in several experimental studies about MID [21].

In Table 2 we show the results according to Entropy, the best results are highlighted in bold.

Table 2. Entropy Results of EFAC vs. the rest of algorithms.

Data sets	KP	KMSF	AD2011	HIMIC	CBMDC	AGKA	EFAC
colic	0.9658	0.9344	0.9503	0.9475	0.9488	0.9451	**0.8978**
dermatology	2.3625	1.731	2.4326	2.4326	2.306	2.4256	**0.6976**
heart-c	0.995	0.996	0.9943	0.9943	0.991	0.9939	**0.9687**
hepatitis	0.7203	**0.5643**	0.7346	0.7346	0.7381	0.7512	0.7344
labor	0.9407	0.7601	0.9348	0.9348	0.9077	0.9456	**0.6486**
lymph	1.7508	1.7856	1.8813	1.703	1.6721	1.7999	**0.9124**
tae	1.5834	1.5818	1.5845	1.5821	**1.5459**	1.5815	1.5515

For clarify these results about the performance of the algorithms, it was necessary find out if they have or haven't significant differences in their performing. For it, we use the methodology recommended by Demsar for the comparison of classifiers in multiple databases [22].

Firstly, we established α =0.05, for a 95% of confidence. Then, for each pair (EFAC vs. Algorithm), we fix the following hypotheses: H0: In the performance of the algorithm EFAC and the other algorithm don't exist significant differences, and H1: In the performance of the algorithm EFAC and that of the other algorithm exist significant differences.

After that, we apply a Wilcoxon test to the results obtained by each pair of algorithms (the algorithm EFAC with each one of the other algorithms). The results of the test are presented in Table 3. Each column show the probability of the Wilcoxon test, and the times the EFAC won, lose or tie with respect other. In each case, if the obtained probability is smaller than 0.05, it is possible to reject the null hypothesis.

Table 3. Wilcoxon test on Entropy.

EFAC vs	Win – Loss – Tie	Probability
KP	7-0-0	**0.018**
KMSF	6-1-0	0.128
AD2011	7-0-0	**0.018**
HIMIC	7-0-0	**0.043**
CBMDC	6-1-0	**0.018**
AGKA	7-0-0	**0.018**

Considering these results we may conclude that the proposed algorithm (EFAC) exceeds in performance to the rest of algorithms except the method KMSF. In this case, the experiment was not enough to say if exist or don't differences between the algorithms.

These results allow saying that metaheuristics based in FA have promissory results for unsupervised classification of mixed and incomplete data. However, it is necessary accomplishing more extensive future experiments.

4 Conclusions

In this paper a new extension to Firefly Algorithm FA is proposed. Our extension, denominated Extended Firefly Algorithm, enlarges the concept of artificial firefly, allowing the modeling of other optimization problems defined in non-numeric domains. The algorithm was applied for the unsupervised classification of mixed and incomplete data. Experimental results allow affirming that the proposed algorithm achieves similar or superior performance respect other methods reported in specialized literature. Therefore we may conclude that the proposed algorithm is able to find the natural structuration of data.

References

1. Handl, J., Knowles, J.: An evolutionary approach for multiobjective clustering. IEEE Transactions on Evolutionary Computation, Vol.11, No.1, pp. 56 – 76 (2007).
2. Ahn, H., Kim, K. J.: Global optimization of case-based reasoning for breast cytology diagnosis. Expert Systems with Applications, Vol.36, pp. 724 - 734 (2009).
3. Prior, A. K. F., Nunes de Castro, L.: The proposal of two bio-inspired algorithms for text clustering. Learning and Nonlinear Models - Revista da Sociedade Brasileira de Redes Neurais (SBRN), Vol.6, No.1, pp. 29 – 43 (2008).
4. Charalambous, C., Cui, S.: A Bio-Inspired distributed clustering algorithm for wireless sensor networks. WIXON'08, ACM (2008).
5. Ahmad, A., Dey, L.: A k-means clustering algorithm for mixed numerical and categorical data. Data & Knowledge Engeneering, Vol.63, pp. 503 – 527 (2007).

6. Hassanzadeh, T., Meybodi, M. R.: A new hybrid approach for data clustering using firefly algorithm and K-means, in Artificial Intelligence and Signal Processing (AISP). 16th CSI International Symposium on. IEEE (2012).
7. Roy, D. K., Sharma, L. K.: Genetic k-means clustering algorithm for mixed numeric and categorical datasets. International Journal of Artificial Intelligence & Applications (IJAIA), Vol.1, No.2 (2010).
8. Errecalde, M. L., Ingaramo, D. A.: A new AntTree-based algorithm for clustering short-text corpora. JCS&T, Vol.10, No.1 (2010).
9. Yang, X. S. Firefly algorithms for multimodal optimization. Lecture Notes in Computer Sciences, pp. 169 – 178 (2009).
10. Lukasik, S., Zak, S.: Firefly Algorithm for Continuous Constrained Optimization Tasks (2009).
11. Wilson, R. D., Martinez, T.R.: Improved Heterogeneous Distance Functions. Journal of Artificial Intelligence Research, Vol.6, pp. 1 – 34 (1997).
12. Azuaje, F.: A cluster validity framework for genome expression data. Bioinformatics, Vol.18, pp. 319 – 320 (2002).
13. Brun, M.: Model-based evaluation of clustering validation measures. Pattern Recognition, Vol.40, pp. 807 – 824 (2007).
14. Huang, Z.: Clustering large data sets with numeric and categorical values. 1rst Pacific - Asia Conference on Knowledge discovery and Data Mining (1997).
15. García-Serrano, J.R., Martínez-Trinidad, J.F.: Extension to c-means algorithm for the use of similarity functions. 3rd European Conference on Principles of Data Mining and Knowledge Discovery, Prague (1999).
16. Ahmad, A., Dey, L.: A k-means type clustering algorithm for subspace clustering of mixed numeric and categorical data. Pattern Recognition Letters, Vol.32, pp. 1062 – 1069 (2011).
17. Ahmed, R.A.: HIMIC: A Hierarchical Mixed Type Data Clustering Algorithm. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.61.6369&rep=rep1&type=pdf (2005).
18. He, Z., Xu, X., Deng, S.: Clustering mixed numeric and categorical data: A cluster ensemble approach. CoRR, http://arix.org/abs/cs/0509011 (2005).
19. Merz, C. J., Murphy, P.M.: UCI Repository of Machine Learning Databases, University of California in Irvine, Department of Information and Computer Science, Irvine (1998).
20. Hsu, C., Chen, C., Su, Y.: Hierarchical clustering of mixed data based on distance hierarchy. Information Sciences, Vol.177: pp. 4474 – 4492 (2007).
21. Villuendas-Rey, Y.: Selecting features and objects for mixed and incomplete data. 13th Iberoamerican Congress in Pattern Recognition. CIARP 2006, LNCS 5197, Springer Heidelberg, La Habana, pp. 381 – 388 (2008).
22. Demsar, J.: Statistical comparison of classifiers over multiple datasets. The Journal of Machine Learning Research, Vol.7, pp. 1 -30 (2006).

Evaluación cuantitativa de la postura humana mediante el análisis digital de imágenes

Jonas Grande Barreto[1], J. Héctor Contreras Angoa[2], Aldrin Barreto Flores[1], Jaime Rebollo Vázquez[2], Ricardo Álvarez González[1], J. Francisco Portillo Robledo[1]

[1] Facultad de Ciencias de la Electrónica, Universidad Autónoma de Puebla.
[2] Escuela de Fisioterapia, Universidad Autónoma de Puebla,
Av. San Claudio y 18 Sur Puebla, México.

Abstract. Postural evaluation is based on traditional methods without a unified approach to obtain quantitative results. Digital image analysis applied to the postural evaluation presents an alternative that complements evaluation and allows obtain better diagnostic. This paper describes an algorithm that applies two methods to obtain quantitative results in postural evaluation. Both methods are designed to identify markers located in specific anatomical parts on the evaluated patient. The first method uses regions analysis for detecting areas corresponding to the markers. The second method uses the Hough transform to detect circles taking advantage of marker's geometry. The results obtained by implementing both methods are the coordinates of the markers, from these points the slope equation is implemented to generate a quantitative evaluation result.

Keywords: postural evaluation, marker identification, image analysis, regions analysis, Hough transform.

Resumen. Las valoraciones posturales se basan en métodos tradicionales sin un parámetro unificado para la obtención de resultados cuantitativos. El análisis digital de imágenes aplicado a la evaluación postural presenta una alternativa que complementa la evaluación y permite obtener un mejor diagnóstico. En el presente trabajo se describe un algoritmo que aplica dos métodos para obtener resultados cuantitativos en la evaluación postural. Ambos métodos fueron diseñados para identificar marcadores ubicados en partes anatómicas claves del paciente evaluado. El primer método utiliza análisis de regiones para la detección de áreas correspondientes a los marcadores. El segundo método utiliza la transformada de Hough para la detección de círculos aprovechando la geometría de los marcadores. Los resultados obtenidos al implementar ambos métodos consisten en las coordenadas de los marcadores, a partir de esos puntos se implementa la ecuación de la pendiente para generar un resultado cuantitativo de la evaluación.

©L. Berthet, J. C. Chimal, E. A. Santos, E. Castillo
Advances in Computer Science
Research in computing Science 81, 2014 pp. 57-65

1. Introducción

La evaluación postural es un campo correspondiente a la fisioterapia [1], está enfocada a la exploración del aparato locomotor principalmente rodillas, cadera y hombros. Se evalúa la posición de las articulaciones, idealmente éstas deben alinearse de tal forma que una línea recta imaginaria cruce a través de ellas. Si la línea no intersecta alguna articulación entonces se diagnóstica una alteración en la postura [1]. El método de evaluación consiste en colocar al paciente detrás de una cuadrícula para ubicar la silueta con puntos de referencia, por lo que la evaluación se basa en la percepción del fisioterapeuta. Los trabajos desarrollados por [2] y [3] mostraron el potencial de utilizar procesamiento de imágenes para analizar el rango de movimiento del cuerpo humano. Otra alternativa que se ha reportado es el uso del sensor KINECT de microsoft, el cual ha tenido aplicación en el campo de análisis biomecánico para rehabilitación como lo muestran los trabajos de [4] y [5], aunque se debe considerar el grado de error del sistema KINECT que se ha reportado por [6].

Para un tema relevante como la salud de un paciente, se desea reducir al mínimo los errores que pueda generar el resultado del procesamiento. Otra técnica que se destaca en el análisis postural es el uso de marcadores colocados por un experto en el cuerpo del paciente como en los trabajos de [7] y [8]. El éxito de sus resultados se debe a que al estar colocados en el cuerpo del paciente se tiene una mayor precisión en la detección y diagnóstico de patologías posturales. En este documento se describe la implementación de dos métodos para la detección de los marcadores.

El primer método consiste en el análisis de regiones para la detección de áreas, en el cual se busca los grupos de píxeles que cumplan un requisito específico y conformen un área determinada. El segundo método implementa la transformada de Hough para la detección de círculos (CHT) [9], esta técnica fue seleccionada con el propósito de aprovechar la geometría de los marcadores. Los resultados que aporta la CHT son las dimensiones geométricas de los marcadores (radio, área) y las coordenadas del elemento detectado. Una vez detectadas dichas agrupaciones se calcula el centroide de los elementos para obtener su ubicación dentro de la imagen. A partir de las coordenadas obtenidas por ambos métodos se implementa la ecuación de la pendiente para obtener un resultado cuantitativo el cual se representa con un valor angular para la posición de las articulaciones.

2. Metodología

Para realizar la captura de las imágenes se desarrolló un sistema de visión por computadora compuesto por una cámara digital, sistema de iluminación y marcadores reflectantes. Un experto en fisioterapia coloca los marcadores a la altura de maléolos exteriores, cóndilos femorales, trocantérs mayor y acrómios. Se capturan imágenes de un fondo al igual que tomas de los planos sagitales izquierdo y derecho del paciente. La primera parte del algoritmo consiste en segmentar la silueta del paciente y generar una imagen que pueda ser procesada mediante los métodos de análisis de regiones y CHT.

2.1. Segmentación de la silueta

El proceso para generar dicha imagen se basa en cuatro pasos: primero se realiza la operación de resta tomando las imágenes de los planos y el fondo para segmentar la silueta del paciente en ambos planos. En el segundo paso se umbralan las imágenes resultantes de la resta a partir de un umbral dinámico como lo reportó [10], la silueta segmentada se presenta como un objeto del cual es posible obtener dimensiones. En el tercer paso se toman las dimensiones del objeto segmentado y se recorta la silueta en las imágenes resultantes en la resta, el recorte se efectúa para reducir el tamaño de la imagen a procesar. En el cuarto paso las imágenes recortadas son sometidas a filtros, gaussiano y mediana, para eliminar el ruido y agrupar los píxeles que conforman el área de los marcadores. La imagen resultante del filtrado proporciona características distintivas a los marcadores pues estos se presentan como *huecos* dentro de la silueta del paciente.

2.2. Análisis de regiones

La implementación del algoritmo inicia con el umbralado de la imagen segmentada, el resultado de este proceso presenta la silueta del paciente con huecos conformados por el área de los marcadores. Posteriormente la imagen es revertida y se aplica una dilatación utilizando un elemento de estructura con dimensiones 3x4 como en el trabajo reportado por [11], la dilatación da a los marcadores un área característica. Posteriormente se etiquetan los objetos presentes en la imagen para analizar sus características geométricas y seleccionar a aquellos que cumplen con la condición requerida, en este caso es el área la cual es calculada mediante la cantidad de píxeles que conforman al objeto.

Para seleccionar los objetos de interés se realiza la comparación de su área, este valor debe encontrase dentro de un rango preestablecido. Los elementos que no cumplen con el requisito son descartados, para los elementos que superan la prueba se calcula el centroide de cada uno. La información de los centroides es ordenada tomando el siguiente orden, el primer elemento es el marcador ubicado en el tobillo (maléolo), seguido por los marcadores de rodilla (cóndilo), cadera (trocánter) y hombro (acrómio).

Las coordenadas de cada marcador se utilizan para calcular la posición angular de las articulaciones mediante la ecuación de la pendiente. El primer valor es calculado a partir de la relación tobillo-rodilla, el resultado es la posición angular de esta ultima. El segundo valor es la relación rodilla-cadera, su resultado es la posición de la cadera. Finalmente la posición del hombro se calcula a partir de la relación cadera-hombro.

2.3. Transformada de Hough

La operación de resta presenta a los marcadores como círculos oscuros, por lo cual la CHT es una buena técnica para identificarlos. A pesar de su gran eficiencia, la CHT tiene algunas complicaciones al detectar círculos con radios menores a 10 píxeles. MATLAB cuenta con una herramienta que implementa la CHT con

un factor de sensibilidad (S), este parámetro mejora las probabilidades para la detección de candidatos que presentan oclusiones o distorsiones en su forma. Sin embargo, al aumentar la sensibilidad sin una referencia pueden aparecer falsos positivos. Para aprovechar la herramienta se diseñó un algoritmo que permite obtener resultados ajustando de forma automática los parámetros de entrada.

Los parámetros de entrada son el radio del candidato a ser detectado y el factor S, los resultados obtenidos del procesamiento son el número y las coordenadas del centro de los círculos detectados. Para verificar que los candidatos detectados sean los marcadores sus centros deben cumplir con la Eq. 1, donde $C(x, y)$ son las coordenada del centro del círculo e I_B es la imagen binaria obtenida del umbralado de la imagen resultante de la resta, la cual es posteriormente revertida con el propósito de presentar a los marcadores como agrupaciones de píxeles con valor lógico de 1. Con este proceso es posible verificar que los centros de los círculos detectados se encuentran dentro de un marcador, los candidatos que no cumplen la condición son descartados.

$$C(x, y) = \{(x, y) \mid I_B(x, y) = 1\} \tag{1}$$

El número de candidatos aprobados debe ser igual al número de marcadores colocados en el paciente ($M = 4$), de no ser igual, la imagen es pasada nuevamente por la CHT pero con un incremento en S(ΔS), los incrementos se realizan hasta cumplir con la condición requerida. Al realizar una retroalimentación se asegura la obtención de resultados deseados. Con las coordenadas obtenidas al aplicar la CHT se utiliza la ecuación de la pendiente para calcular el valor angular de la articulaciones analizadas, la figura 1 muestra el diagrama de bloques del algoritmo.

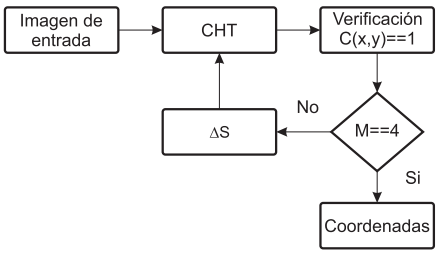


Fig. 1. Diagrama de bloques del algoritmo.

3. Comparación de resultados

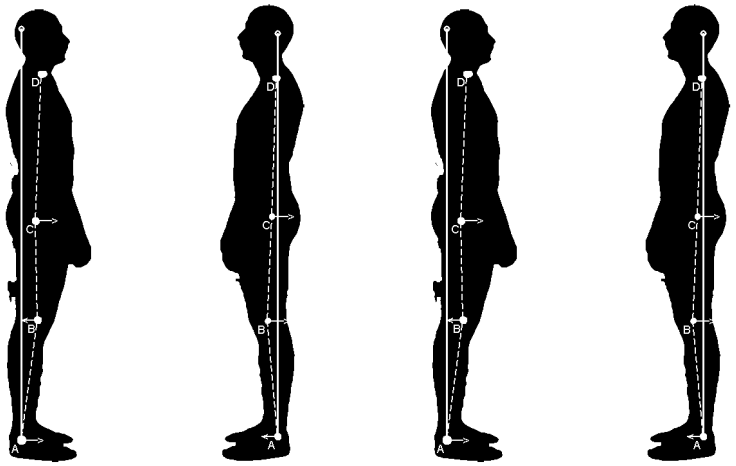
Las pruebas realizadas con ambos métodos se llevaron a cabo en la Escuela de Fisioterapia perteneciente a la Universidad Autónoma de Puebla. Se tomaron fotos de 12 voluntarios a los cuales se les proporcionó dos prendas de vestir y un

gorro, todo en color negro con el objetivo de facilitar la segmentación y detección de los marcadores.

Utilizando el método de análisis de regiones es posible detectar los marcadores, sin embargo esta técnica es susceptible a las variaciones de iluminación lo cual afecta en la segmentación de los pies. Para solucionar ese problema se analizaron las imágenes por segmentos a fin de aislar la region de las extremidades inferiores y aplicar un umbral dinámico a fin de reducir los posibles errores por variación de iluminación. Este método tuvo una efectividad del 80 % con respecto a los 12 pares de imágenes analizadas.

Por el contrario, la CHT es robusta ante las variaciones de iluminación y a las posibles variaciones de tamaño que pueden presentar los marcadores en la imagen segmentada. Esta técnica requiere de menos procesamiento para realizar ajustes y obtener los resultados deseados. Ambos métodos son capaces de detectar los marcadores, sin embargo la CHT tiene una efectividad del 100 % al detectar los marcadores.

Conociendo las coordenadas de los marcadores, la Ecuación de la pendiente es una forma eficiente de obtener valores cuantitativos en la evaluación postural. A partir de las coordenadas es posible trazar líneas que muestren la distribución de las articulaciones, la figura 2 muestra los resultados del análisis sobre los planos sagitales izquierdo y derecho utilizando ambos métodos. La línea recta blanca representa la referencia donde deben ubicarse las articulaciones, la línea punteada representa la distribución de las mismas, las flechas indican la dirección del ángulo, los puntos A, B, C y D son los puntos anatómicos marcados.



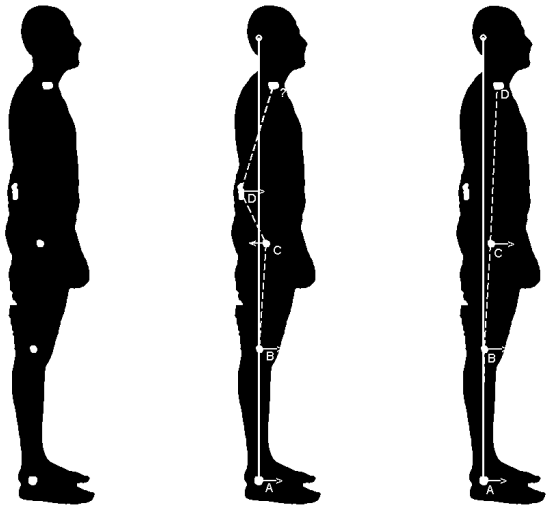
(a) Derecho (b) Izquierdo (c) Derecho (d) Izquierdo

Fig. 2. Análisis de postura: las imágenes a) y b) se obtuvieron con CHT, las imágenes c) y d) se obtuvieron con análisis de regiones.

Tabla 1. Resultados de la valoración postural.

Plano	Tobillo-Rodilla (A)	Rodilla-Cadera (B)	Cadera-Hombro (C)
Izquierdo (Regiones)	85.260°	87.892°	88.135°
Izquierdo (Hough)	84.920°	87.541°	88.144°
Derecho (Regiones)	82.042°	88.522°	87.238°
Derecho (Hough)	82.385°	88.891°	87.749°

La tabla 1 presenta los resultados cuantitativos de la evaluación postural, el diagnóstico terapéutico a partir de la información de la tabla y el apoyo gráfico de las imágenes es una flexión constante en las rodillas y el hombro derecho caído. Estos problemas son causados por la tensión en de los flexores de la cadera para el caso de la rodillas y abducción en la escápula para el caso del hombro. Si la silueta es segmentada por completo sin ninguna alteración, ambos métodos resultan efectivos y tienen muy poca variación en sus resultados. Sin embargo, si la silueta presenta objetos que poseen un área similar a los marcadores, el análisis de regiones presenta problemas para diferenciar entre ambos como lo muestra la figura 3b. Dicho problema puede resolverse agregando más características a evaluar tales como perímetro y redondez de los objetos analizados. Por el contrario, la CHT resulta efectiva en la detección correcta de las regiones de interés debido a los parámetros utilizados para realizar la identificación como lo muestra la figura 3c.



(a) Entrada (b) Regiones (c) Hough

Fig. 3. Comparación de métodos.

La ventaja de la CHT radica en que evalúa los parámetros de forma y dimensiones de los objetos a diferencia del análisis de regiones implementado el cual evalúa únicamente dimensiones haciéndolo susceptible a errores de identificación. La tabla 2 muestra los resultados obtenidos en la valoración de 12 pacientes, la valoración de cada paciente incluye los planos derecho e izquierdo por lo que se analizaron 12 pares de imágenes en cada cada método.

Tabla 2. Resultados de 12 valoraciones.

Método	**Aciertos**	**Errores**	**Origen del error**
Hough	24	0	Ninguno
Regiones	19	5	Incapacidad para distinguir forma del objeto

4. Conclusiones

La falta de higiene postural es uno de los principales problemas en las afecciones del aparato locomotor. Sin embargo las técnicas de valoración existentes carecen de un resultado cuantitativo debido a la falta de herramientas que los proporcionen. El procesamiento digital de imágenes aplicado a la valoración postural presenta una alternativa para la obtención de diagnósticos más completos. Se implementaron dos métodos los cuales aprovecha las características geométricas de los marcadores para la obtención de resultados cuantitativos. Los resultados obtenidos y la forma en que estos se presentan al fisioterapeuta y éste lo hace al paciente, facilitaron el entendimiento, diagnóstico y proceso de rehabilitación que debe seguir el paciente. Este trabajo posee las características para servir como herramienta didáctica al proceso de aprendizaje y capacitación para los estudiantes de fisioterapia. El trabajo futuro a desarrollar incluye el análisis frontal y posterior en la valoración postural al igual que el análisis de movimiento enfocado en la terapia de rehabilitación.

Referencias

1. Peterson-Kendall, F and Kendall-McCreary, E and Geise-Provance, P and Rodgers, M and Anthony Romani, W.: Muscles testing and function with posture and pain, Lippincott Williams & Wilkins, Philadelphia. 2005
2. Iwasawa, Shoichiro, Ohya, Jun, Takahashi, Kazuhiko, Sakaguchi, Tatsumi, Ebihara, Kazuyuki, Morishima, Shigeo.: Human body postures from trinocular camera images. Automatic Face and Gesture Recognition. Proceedings. Fourth IEEE International Conference. Pag, 326–331, 2000, IEEE.
3. Diraco, Giovanni and Leone, Alessandro and Siciliano, Pietro.: Geodesic-based human posture analysis by using a single 3D TOF camera. Industrial Electronics (ISIE), IEEE International Symposium on, Pag, 1329–1334, 2011, IEEE.

4. Kitsunezaki, Naofumi, Adachi, Eijiro, Masuda, Takashi, Mizusawa, Jun-ichi.: Kinect applications for the physical rehabilitation. Medical Measurements and Applications Proceedings (MeMeA), IEEE International Symposium on, Pag, 294–299, 2013, IEEE.
5. Lin, Ting-Yang, Hsieh, Chung-Hung, Lee, Jiann-Der.: A Kinect-Based System for Physical Rehabilitation: Utilizing Tai Chi Exercises to Improve Movement Disorders in Patients with Balance Ability. Modelling Symposium (AMS), 7th Asia. Pag,149–153, 2013, IEEE.
6. García, MJ Gómez and Sisamón, C Castejón, Prada, JC García, Carbone, G, Ceccarelli, M.: Sistemas de captura de movimiento para el caminar humano. Análisis y comparativa humano.
7. Diaz, Christian Andrés, Toro, Maria Luisa, Forero, Johana Carolina, Torres, Andrés.: Detección, Rastreo y Rreconstrucción Tridimencional de Marcadres Pasivos para Análisis de Movimiento Humano. Revista Ingeniería Biomédica. Vol, 3. Pag, 56–68, 2009.
8. D'Amico, Moreno and Bellomo, Rosa G and Saggini, Raoul and Roncoletta, P.: A 3D spine & full skeleton model for multi-sensor biomechanical analysis in posture & gait, Medical Measurements and Applications Proceedings (MeMeA),IEEE International Workshop, Pag, 605–608, 2011, IEEE
9. Atherton, T.J., Kerbyson, D.J.: Size invariant circle detection. Image and Vision computing. Vol, 17, Pag, 795–803, 1999, Elsevier.
10. Bezerra, FN, Paula, IC, Medeiros, FNS, Ushizima, DM, Cintra, LHS.: Morphological segmentation for sagittal plane image analysis, Engineering in Medicine and Biology Society (EMBC), Annual International Conference of the IEEE, Pag, 4773–4776, 2010, IEEE.
11. Arai, Kohei, Andrie, Rosa.: Gender Classification with Human Gait Based on Skeleton Model. Information Technology: New Generations (ITNG), Tenth International Conference, Pag, 113–118, 2013, IEEE.

Identificación automática y segmentación de las hiperintensidades de la materia blanca en imágenes de resonancia magnética

Lizette Johanna Patiño Correa, Oleksiy Pogrebnyak,
Edgardo Manuel Felipe Riverón, Jesus Alberto Martinez Castro

Centro de Investigación en Computación, Instituto Politécnico Nacional, México, D.F, México

Resumen. En este trabajo, se presenta una nueva metodología para la identificación automática y la segmentación de hiperintensidades de la materia blanca que aparecen en las imágenes de resonancia magnética de cortes axiales del cerebro. Para ello, primero se emplea una secuencia de técnicas de procesamiento de imágenes para formar la imagen, donde las hiperintensidades de la materia blanca notablemente difieren del resto de los objetos. Esta etapa de preprocesamiento facilita el proceso posterior de identificación y segmentación de las regiones de hiperintensidades. La metodología propuesta fue probada en 55 imágenes diferentes de resonancia magnética de seis pacientes. Estas imágenes fueron analizadas por el sistema propuesto y las imágenes resultantes de hiperintensidades fueron comparadas con las imágenes manualmente segmentadas por un experto. Los resultados experimentales muestran la tasa promedio de positivos verdaderos de 0.9 y el índice de similitud de 0.7. Al mismo tiempo, los falsos positivos se encuentran en la mayoría de los casos dentro de la materia gris que no causa problemas para el diagnóstico temprano. La metodología propuesta del procesamiento y análisis de imágenes de resonancia magnética puede ser útil para la detección temprana de lesiones de la materia blanca del cerebro.

Palabras clave: imágenes de resonancia magnética, segmentación de imágenes, hiperintensidades de la materia blanca.

Abstract. In this work, a new methodology for automatic identification and segmentation of white matter hyperintensities that may appear in magnetic resonance images of brain axial cuts is presented. For this purpose, firstly a sequence of image processing techniques is employed to form an image where the white matter differ notoriously from the rest of the objects. This pre-processing stage facilitates the posterior process of identification and segmentation of the hyperintensity regions. The proposed methodology was tested on different 55 images magnetic resonance images of six patients. These images were analysed by the proposed system and the resulted images were compared to the images manually segmented by an expert. The experimental results show the mean rate of true positives of 0.9 and the similarity index of 0.7. At the same time, the false positives are found in the most cases within the grey matter that does not cause problems for early diagnosis. The proposed methodology of magnetic resonance image processing and analysis may be useful in early detection of brain white matter deceases.

Key words: magnetic resonance images, image segmentation, white matter hiperintensities

©L. Berthet, J. C. Chimal, E. A. Santos, E. Castillo
Advances in Computer Science
Research in computing Science 81, 2014 pp. 65-74

1 Introducción

Actualmente, las imágenes de resonancia magnética (MRI) son instrumentos importantes usados ampliamente en diferentes aplicaciones médicas. Entre los diferentes tipos de posibles males que se pueden detectar con MRI, las imágenes de cortes axiales del cerebro se usan para detectar varias enfermedades caracterizadas por anomalías de la materia blanca. Las lesiones de la materia blanca se caracterizan por la presencia de hiperintensidades dentro de ella (WMH), las que se pueden encontrar dentro de los tejidos de la materia blanca normal como objetos más brillosos cuando el sistema MRI usa las técnicas de ponderación T-2 y la recuperación de inversión atenuada del líquido (FLAIR) [1].

El problema de la segmentación de la WMH es difícil porque las diferencias en brillo entre regiones normales y dañadas son pequeñas y puede variar en toda la imagen. La segmentación manual es posible, pero es una tarea que consume mucho tiempo y está sujeta a la variabilidad del operador; por ello, la reproducción del resultado de la segmentación manual es difícil; por otra parte, el nivel de confidencia atribuida sufre también [2]. Por estas razones, la segmentación automática de la WMH es preferible, pero es una tarea bastante difícil y pertenece al área de investigaciones activas [3, 4, 5]. Recientemente, un gran número de métodos de segmentación de la WMH fueron propuestos en la literatura. En [2], describen tres generaciones de tales métodos.

En este artículo, presentamos una metodología para la segmentación de WMH completamente automática usando imágenes de escala de grises de ponderación T2 FLAIR. La metodología propuesta puede ser útil para los radiólogos que analizan y diagnostican los males cerebrales sin la necesidad de costosas técnicas basadas en muchos cortes multiespectrales y diferentes ponderaciones que producen un gran volumen de datos para cada paciente. El trabajo está organizado así: después de la Introducción, la sección 2 describe la metodología de segmentación de la WMH; en la sección 3 se presentan los resultados, y las conclusiones se brindan en la sección 4.

2 Metodología de la segmentación de las WMH

La metodología propuesta de la segmentación de las WMH en imágenes MRI consta de varios pasos que se pueden dividir en dos etapas: preparación de la imagen y procesamiento final. En la primera etapa, la imagen es preprocesada para el tratamiento posterior, mediante la eliminación del cráneo y la separación de la materia gris. Una vez que la imagen es preparada, en la etapa final diferentes tipos de tejidos del cerebro se analizan y la materia blanca se separa. Finalmente, las hiperintensidades en la materia blanca se detectan y segmentan. El diagrama de bloques de la metodología propuesta se presenta en la figura 1.

2.1. Etapa de preparación de la imagen

En esta etapa se realizan diferentes técnicas de tratamiento: preprocesamiento, eliminación del cráneo y extracción de la materia gris.

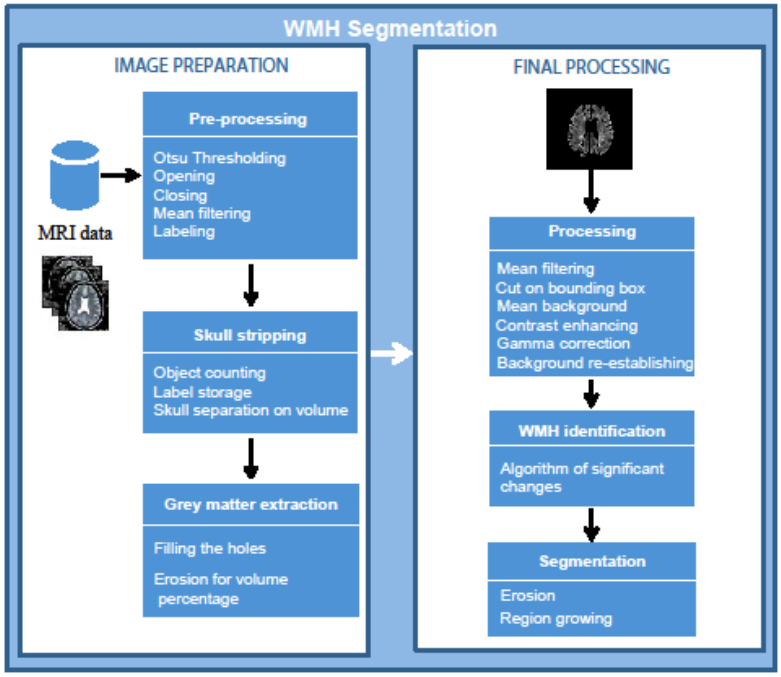


Fig.1. Diagrama de bloques de la metodología propuesta para la segmentación de las WMH.

2.1.1. Preprocesamiento

El paso de preprocesamiento permite disponer de la imagen lista para la eliminación del cráneo. El resultado de este paso es la imagen etiquetada cuyos objetos son el cerebro y partes del cráneo que después se eliminan. El preprocesamiento consiste en el umbralado por el método de Otsu [6], el filtrado morfológico por apertura y clausura, el suavizado con promedio deslizante y el etiquetado de las componentes conectadas.

La técnica de Otsu segmenta los objetos de la imagen que obviamente difieren del fondo, esto es, el cerebro y las cavidades del cráneo en MRI como se muestra en la figura 2. Después del umbralado se aplican tres filtros de mejoramiento de imágenes: filtros de apertura y clausura [7] con el tamaño del elemento de estructura de 4x4 pixeles, y el filtro de promedio deslizante [8] con la ventana de 5x5 pixeles. El objetivo del filtro de apertura es eliminar las uniones indeseables entre objetos después del umbralado Otsu y separar en las imágenes MRI el cráneo y los artefactos que tienen mucho menor tamaño que el cerebro. El filtro de clausura suaviza el contorno del cerebro y rellena hoyos; el filtro de promedio deslizante suaviza la imagen y repone los puntos de las esquinas perdidos en el procesamiento previo.

En el paso siguiente los componentes conectados se etiquetan. Para ello, se empleó el algoritmo de conectividad 8 [9], el cual considera todos los pixeles vecinos tratados que permite etiquetar de una manera más correcta el objeto completo. El objetivo del etiquetado es distinguir los objetos en la imagen por sus etiquetas para poder contarlos y determinar su eventual tamaño. La figura 3 ilustra el etiquetado: en la imagen preprocesada se asigna una etiqueta a cada objeto de diferente nivel de gris.

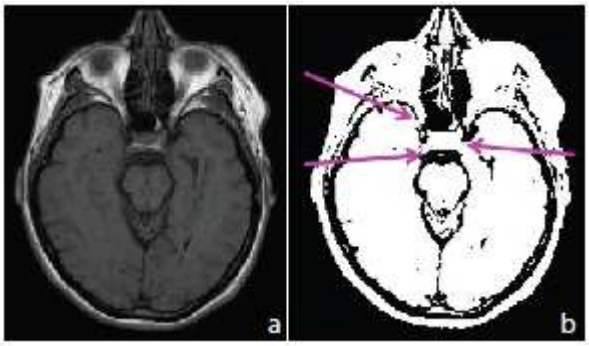


Fig. 2. Umbralado Otsu: a) imagen original; b) imagen resultante del umbralado. Las flechas indican las uniones indeseables entre los objetos.

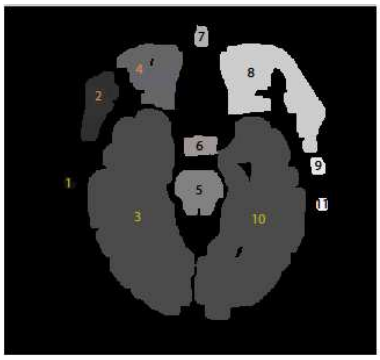


Fig. 3. Componentes conectadas etiquetadas de la imagen del cerebro.

2.1.2 Eliminación del cráneo y separación de la materia gris

La eliminación del cráneo se realiza sobre la base de que el tamaño del cerebro es mucho mayor que el tamaño de los artefactos obtenidos con el preprocesamiento de la imagen. Para ello, primero se cuentan todos los objetos, las etiquetas se guardan y se determina el tamaño de cada objeto. Segundo, los objetos de mayor tamaño se seleccionan y los de menor tamaño se eliminan, separando así los objetos de tejidos del cerebro del resto. Como resultado se obtiene la imagen de los dos lóbulos cerebrales.

En el paso siguiente, solo algunas partes de la materia gris se extraen porque ellas no están bien definidas en la imagen original y sólo se pueden reconocer mejorando el contraste. Ello se ilustra en figura 4a, donde después de mejorar el contraste las regiones de materia gris deben eliminarse ya que representan las hiperintensidades de la materia gris, las cuales se pueden confundir con las hiperintensidades de la materia blanca, resultando la identificación de WMH erróneas.

Para extraer la parte informativa de la materia gris, primero se aplica el algoritmo de relleno de hoyos [7] a las regiones del cerebro porque en algunos casos las WMH se encuentran en los bordes de los hoyos. Después, se aplica el algoritmo de tamaño/porcentaje especialmente diseñado para eliminar un porcentaje de los objetos externos del cerebro.

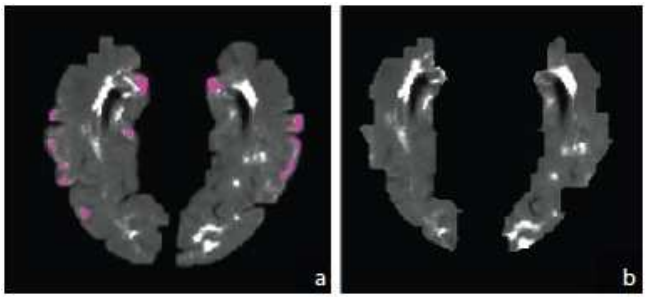


Fig. 4. Imágenes de las regiones de materia gris después del aumento de contraste y la extracción de la parte informativa de la materia gris: a) regiones de contraste aumentado; b) extracción de la parte informativa de la materia gris.

Algoritmo de erosión tamaño/porcentaje

Este algoritmo elimina las regiones usando erosión morfológica [7] con un elemento de estructura de 3 pixeles de diámetro. El algoritmo erosiona la imagen hasta que la diferencia entre la imagen original obtenida después del aumento de contraste de las regiones de materia gris y el relleno de hoyos, $f_h(x,y)$, y la obtenida por erosión, $f_e(x,y)$, se aproxima al porcentaje predeterminado:

1) Calcular el tamaño de los pixeles diferentes de cero V_0 en la imagen $f_h(x,y)$;
2) Realizar la erosión con un elemento de estructura de 3 pixeles de diámetro de la imagen de entrada la cual se designa como $f_e(x,y)$.
3) Calcular el tamaño V_e de los pixeles diferentes de cero en la imagen $f_e(x,y)$.
4) Comparar V_0 y V_e. Si $(V_0 - V_e/V_0)\cdot 100\% < 35.55\%,$ regresar al paso 2.

La imagen obtenida con el algoritmo de tamaño/porcentaje propuesto se intersecta con la imagen binaria del cerebro. Como resultado se obtiene la parte de materia gris (ver la figura 4). La figura 4b muestra que todos los objetos con el contraste mejorado pertenecientes a las hiperintensidades de materia gris se eliminaron.

2.2. Procesamiento final

En esta etapa la imagen se procesa hasta que se obtenga la WMH segmentada. La etapa consta de tres pasos: tratamiento de la imagen, identificación de la WMH y segmentación de la WMH.

2.2.1. Tratamiento de la imagen

El objetivo de este paso es mejorar el contraste de la imagen. Para ello, el filtro de promedio deslizante [8] con la ventana de tamaño 5x5 se aplica a la imagen original del cerebro segmentado. Después la imagen se delimita y se corta en los límites del cerebro usando una caja delimitadora. Seguidamente, el fondo de la imagen se sustituye por el valor promedio de los pixeles diferentes de cero como se muestra en la figura 5.

Después, se mejora el contraste de la imagen mediante el algoritmo de corrección gamma [8] con un factor de 30. Así, se acentúan las lesiones y el fondo de la imagen y sus tamaños se reestablecen. Finalmente, el contraste se mejora agregando la imagen resultante a la imagen original del cerebro segmentado; como se muestra en la figura 6.

La imagen $f(x, y)$ obtenida así es la imagen de entrada para el paso de crecimiento de regiones.

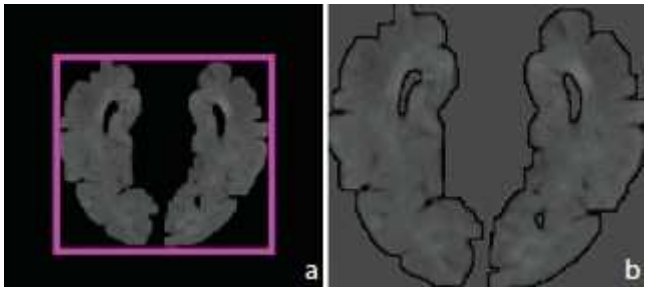


Fig. 5. Tratamiento de la imagen mediante la caja delimitadora y el fondo promedio: a) delimitación de la imagen del cerebro; b) corte de la imagen del cerebro con el fondo promediado.

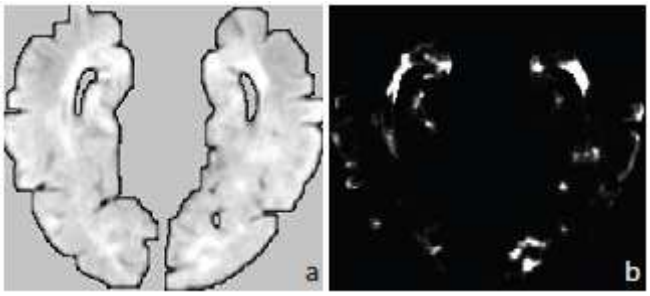


Fig. 6. Mejoramiento del contraste: a) aumento del contraste por saturación; b) la imagen después de aplicar la corrección gamma.

2.2.2. Segmentación e identificación de las WMH

En esta etapa, primero se determinan las regiones de interés con base en los pixeles más brillantes, pues ellos representan las regiones más significativas. Una vez que las regiones de interés se han determinado, el halla centroide de cada región. Los pixeles del centroide resultante son las semillas para el proceso siguiente de crecimiento de regiones; ellas forman la matriz de semillas $S(x, y)$. La figura 7a muestra el resultado del proceso de crecimiento de regiones donde a cada región se asocia sólo un pixel de semilla.

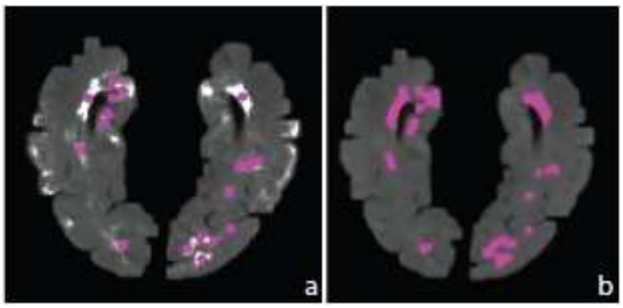


Fig. 7. Segmentación de la WMH: a) detección de las regiones de interés; b) la imagen segmentada.

Los pixeles centroides de las regiones de interés de la imagen resultante del proceso de identificación de la WMH son semillas para la segmentación por crecimiento de las regiones [8, 10]; para este fin se utilizó el algoritmo con conectividad de 8. El resultado obtenido por el procedimiento de crecimiento de regiones es la imagen $g(x,y)$ mostrada en la figura 7b.

3 Discusión de resultados

En esta sección se presentan los resultados de la evaluación del proceso de identificación de las WMH. Para la evaluación se usaron 55 imágenes ponderadas T2 FLAIR de 6 pacientes diferentes con esclerosis múltiple proporcionadas por el Instituto Nacional de Rehabilitación. Estas imágenes son de escala de grises de tamaño 600x600 pixeles. Para la comparación estas imágenes fueron segmentadas manualmente por un radiólogo experto y comparadas con las imágenes segmentadas automáticamente con la metodología propuesta. El criterio empleado para el análisis comparativo fue la diferencia de tamaño: $V_{dif} = \left((V_{man} - V_{aut})/V_{man}\right)\cdot 100\%$, donde V_{aut} es el tamaño de la imagen segmentada automáticamente y V_{man} es el tamaño de la imagen segmentada manualmente.

Las segmentaciones manual y automática fueron comparadas utilizando la tasa de verdaderos positivos (TPR), la tasa de falsos positivos (FPR) y el índice de similitud (SI). La tasa media de los positivos verdaderos (TPR) se define como (Ong et al., 2012): $TPR = V(AS \cap MS)/V(MS)$, donde $V(\cdot)$ es el número de elementos en el conjunto, MS es el conjunto obtenido de la segmentación manual de WMH, AS es el conjunto obtenido de la segmentación automática de WMH y $\cap$ denota el operador de intersección. La tasa media de falsos positivos (FPR) se define como (Ong et al., 2012): $FPR = V(AS \cap \overline{MS})/V(MS)$, donde $\overline{MS}$ es el conjunto de puntos diferentes del conjunto de segmentación de WMH manual. El índice de similitud (SI) se define, como (Ong et al., 2012): $SI = 2V(AS \cap MS)/V(AS) + V(MS)$.

El promedio de verdaderos positivos obtenido es $TPR = 0.9$, lo que significa que la mayoría de las WMH segmentadas manualmente también fueron detectadas automáticamente: 86% de las imágenes tienen TPR en el rango de 0.8 a 1, 7% tienen $TPR = 0.7$, y 7% tienen TPR en el rango de 0.4 a 0.6.

Al mismo tiempo la tasa media de falsos positivos fue $FPR = 0.7$: 50% de las imágenes tienen FPR en el rango de 0 a 0.5, 48% tienen FPR en el rango de 0.8 a 1, y 2% tienen FPR en el rango de 0.6 to 0.7.

El índice promedio de similitud es $SI = 0.7$ que indica una coincidencia significativa entre las dos segmentaciones manual y automática. El 70% de las imágenes tienen SI en el rango de 0.7 a 1 y el 30% restante tienen SI en el rango de 0.3 a 0.6.

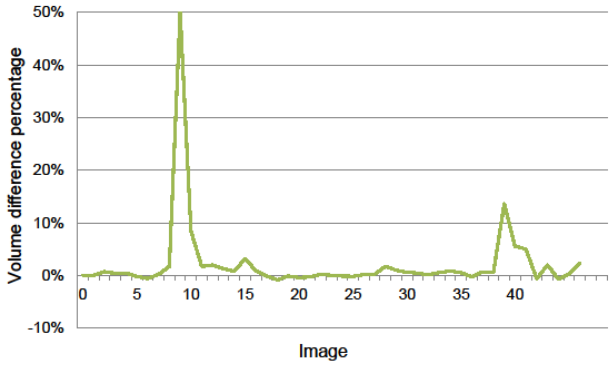


Fig. 8. Diferencia de tamaño entre las segmentaciones manual y automática de las WMH.

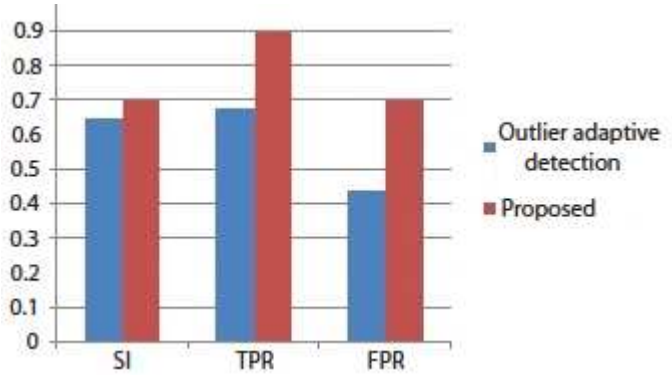


Fig. 9. Gráfica comparativa de los resultados obtenidos por el método de detección automática de valores atípicos y el método propuesto.

Los resultados de la comparación de la segmentación de las WMH en términos de diferencia de tamaño se muestran en la figura 8.

La metodología propuesta fue comparada con el método de detección automática de valores atípicos [3]. Los resultados obtenidos con la metodología propuesta son mejores en términos del índice de similitud y de la tasa de verdaderos positivos, pero existen más falsos positivos al compararla con el modelo de intensidad local [3]. La figura 9 muestra esto gráficamente.

4 Conclusiones

Se ha propuesto una metodología automática para la identificación y segmentación de las hiperintensidades de la materia blanca (WMH) del cerebro. Los resultados finales obtenidos en la identificación y segmentación de las WMH se acercan a los resultados obtenidos con la segmentación de las WMH manual supervisada por un experto. La exactitud de la identificación de las WMH obtenida fue de 87.8%. Las WMH identificadas se segmentan mediante el algoritmo de crecimiento de regiones, obteniéndose una tasa de verdaderos positivos y de índice de similitud de 70%. El trabajo futuro ha de ser orientado hacia la reducción de la tasa de falsos positivos, para

lo que deberán incluirse etapas adicionales de procesamiento de imágenes en la metodología propuesta.

Reconocimientos

Este trabajo ha sido apoyado por el Instituto Politécnico Nacional a través de los proyectos SIP 20140119 y 20141215. Los autores expresan su agradecimiento al Instituto Nacional de Rehabilitación por los datos de MRI brindados.

Referencias

1. Raniga, P., Schmitt, P., Bourgeat, P., Fripp, J., Villemagne, V. L., Rowe, C.C., Salvado, O.: Local intensity model: An outlier detection framework with applications to white matter hyperintensity segmentation. In: Proceedings of 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro 2057 – 2060 (2011)
2. Withey, D.J., Koles, Z.J. Medical image segmentation: Methods and software. In: Proceedings of NFSI-ICFBI Joint meeting of the 6th international symposium on noninvasive functional source imaging of the brain and heart and the international conference on functional biomedical imaging 140 – 143 (2007)
3. Ong, K.H., Ramachadam, D., Mandava, R., Shuaib, I.L. Automatic white matter lesion segmentation using an adaptive outlier detection method. Journal of Magnetic Resonance Imaging 30(6), 807 – 15 (2012)
4. Khademi, A., Venetsanopulos, A.: Robust White Matter Lesion Segmentation in FLAIR MRI. *IEEE Transactions on Biomedical Engineering* 39 (3), 860 – 872 (2012)
5. Samaille, T., Fillon, L., Cuingnet, R., Jouvent, E., Chabriat, H., Dormont, D., Colliot, O., Chupin, M. Contrast-Based Fully Automatic Segmentation of White Matter Hyperintensities: Method and Validation. *PLoS ONE* 7(11): e48953 (2012)
6. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9(1), 62-67 (1979)
7. Soille, P.: *Morphological Image Analysis: Principles and Applications*. Second edition. Springer (2004)
8. Gonzalez, R.C., Woods, R.E.: *Digital image processing*. Third Edition, Pearson, Prentice Hall, (2008)
9. Cheng, C.C., Peng, G.J., Hwang, W.L.: Pixel Connectivity. IEEE Transactions on Image Processing **18** (1), 52 – 59 (2009)
10. Pratt, W.K.: *Digital Image Processing 4th Edition*, John Wiley & Sons, Inc., Los Altos, California (2007)

Communications & Computer Networks

Arquitectura Basada en un Modelo para Votación sobre Medios Móviles

Sandra I. Bautista Rosales[1], Chadwick Carreto Arellano[2] y Eduardo Bustos Farías[2]

[1] Instituto Politécnico Nacional, Escuela Superior de Cómputo, Ciudad de México, México
sibauros@hotmail.com
[2] Instituto Politécnico Nacional, Escuela Superior de Cómputo, Ciudad de México, México
{ccarretoa, ebustosf}@ipn.mx

Resumen. El avance de la tecnología ha llegado a impactar incluso en la democracia, que involucra la toma de decisiones no solo a nivel político, sustentando así lo que se ha denominado como voto electrónico. Actualmente existen países que ya han implementado este tipo de elecciones en base a modelos de votación electrónica que se apegan a la legislación de cada uno de ellos y que además hacen uso de dispositivos especializados para la emisión y recepción del sufragio. Sin embargo, aunque ya existe una cantidad considerable de avances tecnológicos para la captación y conteo de votos, aún no se ha considerado un modelo que pueda basarse en el uso de dispositivos móviles, que le brinde a los electores la oportunidad de emitir su voto desde cualquier lugar en el que se encuentren y con la certeza de que los principios del voto y las garantías procedimentales serán respetados.

Abstract. The advancement of technology has come to impact even in democracy, which involves decisions not only at the political level, thereby supporting what has been called e-voting. Currently there are countries that have already implemented such choices based on models of electronic voting adhere to the laws of each of them and also make use of specialized devices for transmission and reception of suffrage. However, although there is a considerable amount of technological advances to capture and count, still has not been considered a model that can be based on the use of mobile devices, that gives voters the opportunity to cast their vote from wherever they are located and with the certainty that the principles of the vote, and due process will be respected.

Palabras clave: Voto electrónico, Seguridad, Dispositivos móviles, Autenticación, Firma digital.

©L. Berthet, J. C. Chimal, E. A. Santos, E. Castillo
Advances in Computer Science
Research in computing Science 81, 2014 pp. 77-90

I Introducción

Con el avance de las nuevas tecnologías de la información y la comunicación, especialmente el Internet, se ha potencializado el flujo de información a niveles sorprendentes que ha llegado a producir cambios sociales en los países que tienen acceso a este tipo de servicios.

Estos cambios han provocado que la democracia evolucione, los procesos electorales se envuelvan en el flujo de información y la toma de decisiones se facilite; y así todo concurre para considerar la inserción del voto electrónico [2].

Existen países que han ido introduciendo el voto electrónico en sus diferentes procesos electorales de acuerdo con diversos modelos de votación que se apegan al sistema político de cada uno de ellos. Sin embargo, todos estos modelos de votación incluyen dispositivos especializados en la recepción del voto y aún no se ha considerado inclinarse por los dispositivos propios de los usuarios, por ejemplo los dispositivos móviles, para emitir el sufragio.

Por otra parte, el avance continuo de los dispositivos móviles y su capacidad de conectividad con las redes inalámbricas les permite a los usuarios estar comunicados en cualquier momento y desde cualquier lugar (anytime-anywhere); lo que nos lleva a uno de los principales retos que tiene que cumplir toda votación electrónica llevada a cabo por este medio, es decir, proveer seguridad al mandar datos por medio del aire, ya que es uno de los canales de propagación más inseguros para transmitir, por lo que es necesario hacer uso de modelos de seguridad adecuados para cubrir esta necesidad [10]. Esta es la razón por la que el presente trabajo de investigación está enfocado en desarrollar una Arquitectura basada en un Modelo de Votación Electrónica sobre Medios Móviles el cual le permitirá al usuario emitir su voto por medio de un proceso de identificación y autentificación a través de su dispositivo móvil, desde cualquier lugar en el que se encuentre, siempre y cuando cuente con una conexión a Internet.

A continuación en la Sección II se brindará un marco teórico donde se enuncian algunos de los conceptos más importantes para el presente trabajo; en la Sección III se plantea el problema en base a los antecedentes abordados en el marco teórico y se define la problemática a resolver; en la sección IV se describe la Arquitectura propuesta para dar solución a la problemática planteada y finalmente en la sección V se presentan las conclusiones y el trabajo a futuro.

II Marco Teórico

Siendo el voto la expresión pública o secreta de una preferencia ante una o más opciones [1], se puede decir que el voto electrónico (VE) tiene el mismo objetivo, con la excepción de que se emplean mecanismos electrónicos para realizarlo por lo que se pueden identificar dos aspectos del mismo [2]:

a) VE en *sentido amplio*: es todo mecanismo de elección en el que se utilicen los medios electrónicos o cualquier tecnología en las distintas etapas del proceso electoral, principalmente el acto efectivo de votar.
b) VE en *sentido estricto*: es el acto preciso en el que el emitente del voto deposita o expresa su voluntad a través de medios electrónicos (urnas electrónicas) o cualquier otra tecnología de recepción de sufragio.

Teniendo en cuenta esta clasificación, el presente trabajo considerará el voto electrónico en el sentido amplio, lo que lleva a definir algunos conceptos básicos para este aspecto.

Existen diversos sistemas de voto electrónico que actualmente ya son utilizados en diferentes países, principalmente los occidentales; de manera general se identifican tres tipos de sistemas principales y un cuarto [2] que va en proceso de adoptarse por ser más reciente. Los sistemas antes mencionados son:

1) Mediante tarjeta perforada. El elector recibe una tarjeta, en la cual debe perforar su opción por medio de un aparato. Este sistema es un tanto problemático, pues la precisión de la perforación depende del usuario y podrían no contarse adecuadamente la perforación de cada tarjeta. Este es un sistema obsoleto, pero aún lo continúan empleando.
2) Mediante un aparato lector. El votante realiza marcas con un bolígrafo en una papeleta, para que después éstas puedan ser introducidas en un aparato lector y se cuente el voto. El votante no entra en contacto directo con la tecnología, pero sí su papeleta.
3) Mediante aparatos de grabación directa. Usan aparatos similares a un cajero automático, el elector establece su preferencia por medio de una pantalla táctil o un teclado. Puede que el mismo aparato registre el voto o que el voto se grabe en un soporte externo. Tras emitir su voto, el votante utiliza su tarjeta a modo de una papeleta tradicional, introduciéndola en una urna, que a su vez será un aparato lector de tarjetas magnéticas, y que realizará el recuento.
4) Voto electrónico remoto. Es el que provee que el votante no deba desplazarse hasta el colegio electoral y pueda emitir su voto a través de la red (puede ser interna o desde cualquier plataforma conectada a Internet).

El modelo sobre el cual se basa la Arquitectura propuesta corresponde a éste último tipo de sistema..

En general la ***seguridad***, se refiere al cumplimiento de los servicios de seguridad que se listan a continuación [3]:

- **Confidencialidad**. Garantiza que la información privada pueda ser accedida únicamente por las entidades autorizadas.

- **Integridad de los datos**. Se refiere a la protección de los datos de tal manera que no puedan ser modificados, destruidos o extraviados de una manera maliciosa o accidental.
- **Autenticación**. Es un servicio relacionado con la identificación de identidades o de datos. La autenticación de una entidad es la confirmación de su identidad, es decir, una comprobación de que es quien dice ser. La autenticación de datos se refiere a la validez de los datos, lo que implica la integridad de los mismos.

- **No rechazo**. Asegura que el remitente de cierta información no pueda rechazar/negar su transmisión o contenido y que el receptor no pueda negar su recepción o contenido.

La ***firma digital*** es un esquema matemático que garantiza la privacidad de la conversación, la integridad de los datos, la autenticidad del mensaje/emisor digital y el no repudio del remitente [4]. También se puede decir que es un método criptográfico que asocia una identidad, ya sea de una persona en particular o de un equipo a un mensaje enviado a través de transmisión por la red.

La ***firma a ciegas*** es una clase de firma digital producida para mensajes que se mantienen ocultos para el signatario y por lo tanto una tercera entidad solicita la firma [5]. Este concepto fue introducido por Chaum con la finalidad de garantizar anonimato en los sistemas de pago electrónico [6], pero actualmente también se le ha dado uso en los sistemas de votación electrónica. Las firmas a ciegas requieren de dos entidades, el usuario quien solicita la firma y el signatario quien firma el mensaje.

De manera general, un sistema de votación electrónica por Internet debe cubrir con todos los requisitos funcionales del proceso electoral, así como los servicios de seguridad necesarios para protegerse de ataques potenciales provenientes de la red. Algunos de los requisitos esenciales son los siguientes [5]:

- *Autenticación*: sólo los votantes incluidos en el padrón electoral serán autorizados para emitir su voto.

- *Anonimato y no coerción*: nadie debe ser capaz de determinar el valor del voto ni de vincularlo con el votante.
- *Unicidad*: ningún votante debe votar más de una sola vez.
- *Verificación y auditoría*: debe ser posible verificar que al final del proceso electoral, todos los votos fueron contados correctamente.

Otros requisitos propuestos por [7] son:

- *Precisión*: Se debe evitar que el voto emitido sea alterado, duplicado o eliminado por cualquier persona. Cada voto legítimo debe ser contabilizado correctamente.

- *Simplicidad*: El proceso de votación debe ser lo más simple posible. En otras palabras, una interfaz de elección electrónica amigable al usuario y que no necesite aprender técnicas complejas y cualquier equipo adicional.

III Planteamiento del Problema

Los avances tecnológicos aunados a las técnicas criptográficas han sido un factor para la paulatina inserción del voto electrónico dentro de los sistemas tradicionales de sufragio. El Internet y diversos dispositivos electrónicos facilitan la captura, recepción y transmisión del voto, ofreciendo así un proceso más cómodo y eficiente para todos los participantes.

No obstante, el nuevo sistema de votación que se deriva del voto electrónico, es decir, el e-voting remoto acarrea también ciertos problemas o riesgos adicionales a los que ya existen en el sistema tradicional. Por ejemplo, la autenticación del usuario, ya que debe de ser identificado para tener los respectivos permisos para votar y a su vez guardar su anonimato al momento de emitir su voto.

Otro problema viene de considerar el Internet como el principal medio de transmisión, pues es necesario implementar los esquemas de seguridad para votaciones electrónicas que ofrezcan seguridad al sistema y confidencialidad a los votantes.

Ya que generalmente el voto es capturado electrónicamente y enviado a una urna electrónica para ser contabilizado al finalizar el periodo de votación, la auditoria se convierte en otro problema, ya que no hay una garantía de que esos votos que van siendo almacenados no sean alterados mientras finaliza la elección.

En la actualidad hay muchos protocolos y esquemas de votación electrónica que van de acuerdo a ciertos modelos de votación. Por ejemplo, "Un esquema verificable del voto electrónico por Internet" [7] proporciona un esbozo de cómo dar seguimiento al proceso de la votación, especificando 4 fases, a saber, fase de registro, autenticación, votación y conteo, además propone cinco entidades participantes: Votantes, Autoridad Certificadora, Centro de autenticación, Centro de supervisión y Centro de conteo, siendo la mayor aportación de este trabajo la verificabilidad.

Otro trabajo titulado "Un Protocolo Seguro de Voto Electrónico para las Elecciones Generales" [8] propone un centro de voto que está compuesto por cuatro pasos con los que navega a través de cinco módulos.

Un tercer protocolo es "Un Mecanismo de Votación Anónima basado en el Protocolo de Intercambio de Claves" [9], éste hace uso de la firma a ciegas y del protocolo de Diffie-Hellman para el intercambio de llaves entre el votante autenticado y el servidor encargado de recibir el mensaje correspondiente al voto. Una de las desventajas de este protocolo es que no permite al votante verificar que su voto haya sido contado correctamente.

En la tabla I podemos observar la comparativa de las diferentes propuestas de modelos y protocolos de votación.

TABLA I.
COMPARATIVA DE POTOCOLOS DE VOTACIÓN ELECTRÓNICA

Protocolos				
Requerimientos	**Protocolo de Liaw [8]**	**Protocolo de Chang-Lee [9]**	**Protocolo de Chun-Ta Li [7]**	**Modelo y Arquitectura propuestos**
Precisión	✓	✓	✓	✓
Simplicidad	X	✓	✓	✓
Apegado a la Ley	✓	✓	✓	✓
Verificabilidad	✓	X	✓	✓
Privacidad	✓	✓	✓	✓
Esquema sobre Internet	✓	✓	✓	✓
Basado en redes inalámbricas	No especifica	No especifica	No especifica	✓
Contemplan dispositivos móviles	X	X	X	✓

IV Arquitectura Propuesta

Con los modelos de votación electrónica que se han propuesto hasta el momento, se puede observar que la mayoría toma en cuenta lugares específicos para que el elector acuda a emitir su voto y/o contemplan dispositivos dedicados a la recepción y conteo del voto. Es verdad que ya existen modelos planteados para un esquema basado en Internet, pero sus descripciones en cuanto al tipo de comunicación que emplearan son muy vagas o casi nulas y dichos modelos se enfocan principalmente en el acto efectivo de votar, dejando de lado las especificaciones y protocolos que se deben de seguir al hacer uso de una red alámbrica o inalámbrica.

En cuanto a la seguridad, la mayor parte de los protocolos propuestos usan relativamente mecanismos sencillos de autenticación, por ejemplo, tarjetas inteligentes y/o una comprobación de que el votante se ha registrado previamente en una base de datos, por lo cual es necesario hacer uso de distintos mecanismos de autenticación para poder brindar confiabilidad en el sistema de votación.

Por otra parte, los modelos de votación electrónica ya propuestos solamente mencionan que las diferentes entidades gubernamentales deben participar y vigilar que el proceso de la votación se realice correctamente, pero dichos modelos no contienen un apartado dedicado a la parte legislativa que es aplicada en cada una de las naciones democráticas.

La importancia de llevar la votación electrónica a un esquema móvil es de vital importancia, ya que brindaría a los votantes un modelo de votación mucho más conveniente para ellos, haciendo uso de las bondades que brindan sus propios dispositivos móviles, es decir, permitirles emitir su voto desde cualquier lugar en dónde se encuentren (y esté disponible una red inalámbrica), sin necesidad de trasladarse a un lugar donde se encuentre algún dispositivos especializado en recibir o contabilizar el voto.

El modelo propuesto es una alternativa que se podría ir adoptando en las diversos países democráticos, ya que además de brindarle beneficios a los votantes al romper con las barreras geográficas, también es conveniente para las entidades encargadas de gestionar todo el proceso de la votación, pues evitaría gastos innecesarios en material (papel, bolígrafos e incluso dispositivos especializados que pueden llegar a ser muy costosos) y la mayor parte de los esfuerzos se podrían enfocar en garantizar la seguridad del sistema y del votante.

El Modelo de Votación Electrónica Móvil propuesto, consta de cuatro capas como se observa en la Fig. 1

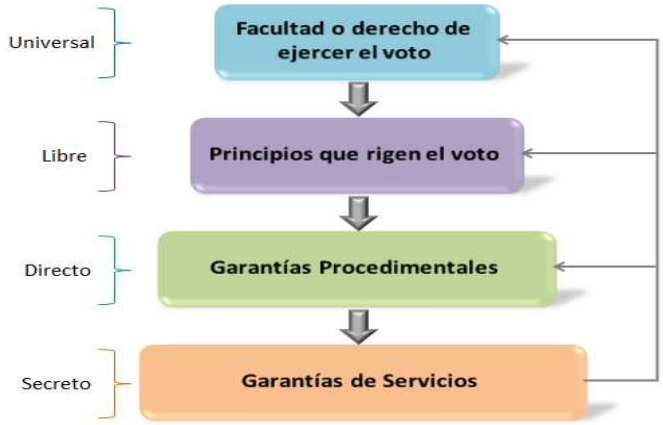


Fig. 1. Modelo propuesto de Votación Electrónica Móvil (Fuente propia)

Cada capa está enfocada a diversos aspectos y fases dentro del proceso de votación, a continuación se describe cada una:

A. ***Facultad o derecho de ejercer el voto***: Se refiere primeramente al proceso de registrar a los votantes que cumplen con los requisitos básicos para poder votar, de esa manera se les otorgará un nombre de usuario y contraseña para después hacer uso de ellos en el proceso de identificación del votante.
B. ***Principios que rigen el voto***: En esta capa se especifican el proceso de autenticación del votante, la cual se realizará por medio de la Firma Electrónica y la implementación del algoritmo de encriptación RSA para autenticarse a través de su dispositivo móvil, además de especificar los principios del voto, como son el sufragio universal, libre, igual y secreto, por mencionar algunos.

C. ***Garantías Procedimentales***: Esta capa involucra la emisión de documentos, certificados, por ejemplo el certificado de autenticación de usuario; también se especifican los aspectos legislativos que regirán el proceso de la votación y se definen las reglas de acuerdo con la legislación del país o de la entidad organizadora de la votación.

D. ***Garantías de Servicios:*** En esta capa se especifican las garantías con las que debe cumplir cada procedimiento especificado dentro de la votación, por ejemplo la transparencia, la verificación de la autenticación, verificación del voto emitido por medio de un certificado digital, la integridad de la información, comunicación segura en el envío de datos a través de la implementación del algoritmo AES, entre otros.

Cada capa tratará de cubrir una de las cuatro características esenciales del voto tradicional que también se deben aplicar al Voto Electrónico sobre medios móviles, como se enlistan a continuación:

- ***Universal***, porque el derecho al voto es inherente a todos los ciudadanos, sin más restricción que cumplir las condiciones que establece la entidad que organiza.
- ***Libre***, porque el ciudadano vota por la opción de su preferencia de manera consciente e informada, sin que persona alguna ejerza presión o lo condicione de alguna manera.
- ***Secreto***, porque el ciudadano emite su voto privado, y ninguna persona puede requerirle información sobre el sentido del mismo.
- ***Directo***, porque el ciudadano acude a votar personalmente.

Para verificar la funcionalidad del Modelo de Votación Electrónica, se ha diseñado una Arquitectura para poder montar un caso de estudio que consiste en una Urna Electrónica Móvil que será implementada sobre dicha Arquitectura que se muestra en la Fig. 2, donde también se aprecia en qué parte de la Arquitectura se implantará cada capa del Modelo propuesto.

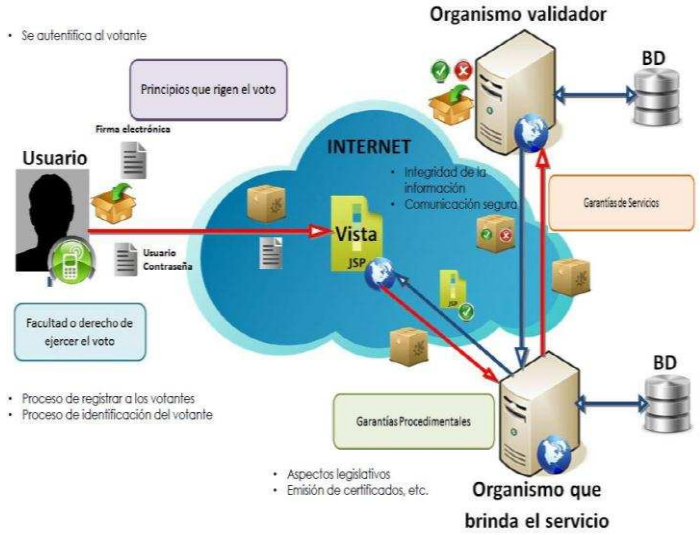


Fig. 2. Descripción de cada capa del Modelo en la Arquitectura propuesta. (Fuente propia)

El flujo de datos que seguirá la Arquitectura será como se describe a continuación y se aprecia gráficamente en la Figura anterior

- El usuario ingresa a la interfaz de la aplicación, si aún no está registrado no podrá acceder a ningún servicio y tendrá que proceder a registrarse.

- Para el registro, el usuario tendrá que mandar un paquete de datos a una vista a través de la red, es decir, debe seleccionar la opción de registrarse, llenar los datos que se le pidan y enviar ese paquete de datos a través de Internet.
- Estos datos serán redireccionados a un servidor, que es el organismo que brinda el servicio con la aplicación de voto electrónico móvil.
- El organismo que brinda la aplicación de voto envía el paquete de datos a otro servidor, que le pertenece al organismo validador de registro, en este caso un administrador se hará cargo de verificar que los registros se realicen de manera correcta.
- El organismo validador del registro revisará los datos del paquete y regresará una respuesta de éxito o fallo en el Registro.
- Dependiendo de la respuesta del organismo validador, será la acción que emita el organismo que brinda la aplicación de votación electrónica móvil y por ende las vistas que se le desplieguen al usuario.

V Caso de Estudio

Las tecnologías que se utilizaron para desarrollar el Caso de Estudio se dividen en tres módulos y se describen a continuación:

1. **Aplicación Móvil**
    a. Sistema Operativo Móvil: Android
    b. Lenguajes de Programación: SDK 4.0 Ice-Creame Sandwich API level 15 debido a las herramientas de seguridad, criptografía y canales seguros de comunicación que se han incluido a partir de esta versión.
    c. Sistema Gestor de Base de Datos: SQLite ya que es el sistema gestor de base de datos que ocupa android y brinda la facilidad de conectarse con MySQL.
2. **Aplicación Web**
    a. Lenguajes de programación
        i. Java SE7U25 JDK: Se eligió este lenguaje debido a la portabilidad que brinda la máquina virtual en diversos sistemas operativos.
        ii. JSF 2.0: (Java Server Faces) Este framework simplifica el desarrollo de interfaces web y permite el uso del modelo vista controlador.

iii. Primefaces 4.0: Es un complemento de código abierto para JSF que ayuda a mejorar la interacción del usuario con la página web.

3. Servicio Web

a. Lenguajes de programación: Java para poder programar el algoritmo de cifrado AES y RSA ya que es el lenguaje provee de algunas librerías para la implementación de estos algoritmos.
b. Sistema Gestor de Base de Datos: MySQL Server porque debido a la compatibilidad que existe entre Java Server Faces y los conectores brindados por Java se facilita la conexión entre estas dos herramientas.

Restricciones del Sistema

- El sistema estará enfocado a smartphones con el sistema operativo android con una versión igual o superior a la 4.0.
- El sistema debe contar con una conexión a internet mediante datos móviles para poder llevar a cabo la votación.
- El sistema deberá de contar con un dispositivo de almacenamiento externo que contenga la información cifrada del usuario.

El sistema requerirá de la validación inicial de usuarios por parte del órgano organizador o regulador del evento.

Cabe aclarar que la inserción del voto electrónico a través de dispositivos móviles debe ser una transición paulatina, usando primeramente este tipo de votación como un apoyo a la votación tradicional mientras se sigue trabajando en todos los elementos que tienen que converger para que la votación sobre medios móviles se pueda implementar en su totalidad.

VI Conclusiones y trabajo a futuro

En el presente trabajo se describió la propuesta y avance de una Arquitectura basada en Modelo de Votación Electrónica sobre Dispositivos Móviles. Como se pudo observar, el Modelo se basa en cuatro capas para cubrir los diferentes aspectos que se deben de considerar en el Voto Electrónico y la Arquitectura nos brinda la conectividad entre cada uno de los componentes. Pretende como aportación el poder

garantizar no solamente la identificación y autenticación del votante, sino también pretende lograr la protección de la identidad del usuario, así como la seguridad de los votos emitidos. Por otro lado, con la implementación del Modelo en la Arquitectura se brindará transparencia a lo largo de todo el proceso y no solamente al finalizar la votación.

Actualmente se está trabajando en pruebas de seguridad del mecanismo de Autenticación de usuarios de la Urna Electrónica Móvil (caso de estudio) y pruebas al canal por donde fluyen los datos, el cual es una parte de un Modelo de Seguridad en Redes Móviles que la firma electrónica del usuario para su funcionamiento [12] y en la aplicación para el usuario en el sistema operativo Android.

Agradecimientos

Los autores del presente trabajo agradecen a ESCOM, CIC, ESIME, COFAA y a la SIP por las facilidades proporcionadas para el desarrollo del presente trabajo.

Referencias

1. Real Academia Española, Diccionario de la lengua española, 22ª ed. 2010, http://lema.rae.es/drae/
2. Téllez Valdés, Julio. El voto electrónico, México, Tribunal Electoral del Poder Judicial de la Federación (2010) http://www.te.gob.mx/documentacion/publicaciones/Temas_selectos/14_voto.pdf
3. F. Rodríguez-Hernríquez, N.A. Saquip, A. Díaz Pérez, C. K. Koc. Cryptographic Algorithms on Reconfigurable Hardware (Signals and Comminication Technology). Springer Verlag, New York, Inc. (2006)
4. Kaur, R. Digital Signature. Guru Nanak Dev University, India. International Conference on Computing Sciences (2012)
5. M. de L. López García. Diseño de un protocolo para votaciones electrónicas basado en firmas a ciegas definidas sobre emparejamientos bilineales. Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional. México (2011)
6. D. Chaum, R. L. Rivest and A. Sherman. Crypto 82 in Advances in Cryptology 1981 to 1997, Vol. 1440 of Lecure Notes in Computer Science, p. 13-19. Springer Berlin (1999)
7. Chun-Ta Li, Min-Shiang Hwang, Yan-Chi Lai. A verifiableElectronic Voting Scheme over the Internet. Sixth International Conference on Information Technology: New Generations, Las Vegas Nevada (2009)
8. H. T. Liaw. A secure electronic voting protocol for general elections. Computers & Security (2004)
9. C.C. Chang and J. S. Lee. An anonymous voting mechanism based on the key exchange protocol. Computers & Security (2006)
10. Mudhakar Srivatsa. Who is Listening? Security in Wireless Networks. IEEE-International Conference on Signal processing, Communications and Networking, Madras Institute of Technology, Anna University Chennai, India, pp. 167-172 (2008)

11. Gentles, Donovan and Sankaranarayanan. Biometric Secured Mobile Voting. Kingston, Jamaica (2011)
12. Hernández O. Luis E. Modelo de Seguridad para Redes Aplicado a Dispositivos Móviles. Escuela Superior de Cómputo, Instituto Politécnico Nacional (2013)
13. L. Bass, P. Clements, R. Kazman, Software Architecture in Practice, 2nd Edition, Addison Wesley (2003)

Databases & Software Technology

A Configurable Web System for Generating E-Commerce Web Sites

Carlos R. Jaimez-González, René Ortíz-González

Departamento de Tecnologías de la Información,
Universidad Autónoma Metropolitana, Unidad Cuajimalpa,
Av. Vasco de Quiroga 4871, Col. Santa Fe Cuajimalpa, C.P. 05300, México, D.F.
{cjaimez,209363198}@correo.cua.uam.mx

Abstract. In order to expand their opportunities of selling products and services, conventional stores have been creating their own e-commerce web sites, where they can carry out electronic transactions from any place at any time. Store owners and their clients benefit from e-commerce web sites, because they are permanently online and provide catalogues of products and services with images, descriptions, specifications, prices, comments from clients, recommendations, among other features; so clients have enough information to make a decision on the products and services they will buy. There are some tools available on the web that generate this type of web sites, but they do not provide the functionality of displaying the web site on mobile devices automatically. This paper presents a configurable web system for generating and administering e-commerce web sites, which are automatically displayed on mobile devices, using the responsive web design pattern.

Keywords. Configurable Web System; E-Commerce Web Site; Responsive Web Design; Web-Based Application.

1 Introduction

In order to expand their opportunities of selling products and services, conventional stores have been creating their own e-commerce web sites, where they can carry out electronic transactions from any place at any time. Store owners and their clients benefit from e-commerce web sites, because they are permanently online and provide catalogues of products and services with images, descriptions, specifications, prices, comments from clients, recommendations, among other features; so clients have enough information to make a decision on the products and services they will buy. A fundamental feature of an e-commerce web site is the shopping cart, in which a client puts the products wanted, so that when the selection is finished, clients proceed to checkout, where they pay for the products through credit card, debit card, paypal, or other mechanism. Once the products have been paid for, in most of the cases the seller sends them to the address provided by clients.

Some of the main features and components of e-commerce web sites are the following: user registration for clients; visualization of catalogues of products and ser-

©L. Berthet, J. C. Chimal, E. A. Santos, E. Castillo
Advances in Computer Science
Research in computing Science 81, 2014 pp. 93-103

vices with several levels of categories; visualization of information of each product or service, such as its name, description or specification, price, and others; product finder; recommendation of products; shopping cart for storing products; selection of payment and online payment; selection of delivery for products; visualization of purchasing history; modules for administration of catalogues, such as categories, products, payments, deliveries, sales history, sales statistics, among others.

The use of devices such as smartphones and tablets has increased the development of mobile applications, which allow accessing different web sites. The visualization of a web site in a mobile device is different to the one in a laptop or desktop computer; this is mainly due to the difference in the resolution of mobile devices. In order to have an effective and correct visualization on any device or computer, it is necessary in most of the cases to modify the web site, so that it can be displayed in a satisfactory way according to the device. Although there are some tools that generate e-commerce web sites, such as *osCommerce* [1], *Magento* [2], *Prestashop* [3], *Opencart* [4], none of them provide the functionality of having the resulting web site visualized automatically on mobile devices. This paper presents a configurable web system for generating and administering e-commerce web sites, through the selection of some configuration options. Additionally, the e-commerce web sites created will be displayed automatically on mobile devices, without requiring the user to modify the web site.

The rest of the paper is organized as follows. Section 2 presents the main components of an e-commerce web site and some related work. The architecture of the web system is introduced in section 3. Section 4 describes the implementation of the web system, including a site map, some web user interfaces, and the technologies used to implement it. Finally, conclusions and future work are provided in section 5.

2 Related Work

This section provides a description of the main components of an e-commerce web site, and reviews some existing tools that generate e-commerce web sites.

2.1 E-Commerce Web Site Components

There are three components of an e-commerce web site: online store or front-end; administration panel or back-end; and the data persistence component. The main features of each component are provided in the following paragraphs.

Online store. The client interacts with this component, in which it is possible to visualize the offered products, carry out transactions, and there is communication with the store. The main features and modules that compose it are the following: products catalogue, product view, categories of products, promotions, recommendations, product search, user registration, shopping cart, payments, deliveries, among others.

Administration panel. This component is used by the system administrator to create the online store; it has options to add, update and delete catalogues in the system, such as products, categories, clients, payments, deliveries, promotions, etc. It can also search historical data, as well as statistics of the store, and managing the entire

system. The main features and modules that compose it are the following: management of products, categories, payments, deliveries, statistics, historical data, invoices, promotions, users, communication between users, among others.

Data persistence. This component is in charge of storing the information of the system; it guarantees its integrity and persistence. Most of the e-commerce systems examined are based on relational models of persistence, among which are the following database management systems: MySQL, PostgreSQL, Oracle, SQL Server.

2.2 Existing Tools

osCommerce [1]. According to its documentation, it is an online program for e-commerce and administration, which was developed in PHP by Harold Ponce de Leon. In order to create an online store with *osCommerce*, it is required a MySQL database, and an Apache web server. *osCommerce* is composed of two parts: the products catalogue; and the administration module to manage the online store; it also allows updating products, categories, languages, new offers, currency types, orders, clients, payment and delivery types, etc. An online store created with *osCommerce* is divided in three panels: the left panel shows categories of products, selection of manufacturers of products, product finder with advanced search, new products, information for shipping, private notice, conditions of use, and a contact form; the central panel shows the catalogue of products with images, descriptions, and prices; and the right panel shows the shopping cart, the bestsellers, special products, product reviews, and conversion of currencies. At the top of the web site there are three buttons: one for accessing the shopping cart content, one for going to checkout, and one for the user account.

Magento [2]. It is a system for creating online stores, where its source code is available and free. It is possible to upgrade to commercial versions, and increase the functionality of the online store. The main features of *Magento* are the following: it can create traditional online stores with physical products; but it also allows creating online stores with downloadable material, such as music or ebooks. It allows multi-stores; it has a search system; it allows complete customization of the web site through an administration panel; it has a module for orders management; etc. An online store generated with *Magento* is divided in three panels: the left panel shows two random offers of products, and popular tags for searching products; the central panel shows some other offers of products with discounts, and some of the best seller products with their images, descriptions, and prices; and the right panel shows a module for comparing products, the shopping cart, an advertisement, and a survey. At the top of the web page, there are the following links for the user of the online store: my account, my wish list, my shopping cart, checkout, log in, and product finder.

Prestashop [3]. It is an open source solution that combines the PHP language, an Apache web server, and a MySQL database. It allows store owners to maintain their online stores without any license costs. *Prestashop* is flexible in terms of design and distribution of the web page. It uses a hook mechanism to put any module in any position on a web page, such as the shopping cart, currencies, product search, etc. The hooks are calls to components, which are positioned using JavaScript code. Among its main features are the following: multi-language support; it is possible to create cus-

tomized web pages, including images and other multimedia components; several payment modules; integrated clients module; definition of tangible and virtual (digital) products; use of bar codes, etc. An online store generated with *Prestashop* has a three-column layout, where the central column has the catalogue of the store, and a list of outstanding products. The left column contains a section with tags of products, categories, manufacturers, and general information, such as deliveries, legal notice, conditions of use and contact. The right column has the shopping cart at the top, and a section for new products. The user menu is displayed at the top of the site, with links for contact, site map, favorites, currency converter, user account, and product search.

Opencart [4]. It is an open source e-commerce system, which is based on the PHP programming language. *OpenCart* was created with the following features: it is a light version derived from *osCommerce*; it has good documentation for developers; interchangeable template system with easy development; it uses the Model-View-Controller (MVC) design pattern; it is modular; multiple languages; several ways of payments; comments on products; different ways of delivery; evaluation of products with star ratings; etc. An online store generated with *OpenCart* is simple and orientated to the Web 2.0. Its interface preserves the layout of three columns: on the left column the categories of products, and other information, such as about us, privacy policy, terms and conditions, contact us, and site map; the catalogue of products is located on the center, with images, descriptions, and prices for each product; the right column contains the shopping cart and the best sellers. The user menu in this system is in the form of tabs, and it has the main page, log in, account, shopping cart, checkout, product search by product, by category, and advanced search.

Other systems. There are several systems for creating e-commerce web sites. The most used systems were described and analyzed in the previous sections, but some other systems can be found in the comparative table of shopping cart software shown in [5]. However, there are many of them unavailable or with obsolete versions. Some of the systems still in use are the following: Zen-Cart [6], Go-eCommerce [7], Shoppizer [8], Konakart [9], and JadaSite [10]. The scripts described in this section have been created using Java Server Pages (JSP) mainly. These systems have not been used as much as the other systems analyzed, because they are complex Java applications; they are difficult of customization; they do not have a version for displaying the resulting e-commerce web sites on mobile devices; and given its architecture and implementation, it is difficult to incorporate such functionality.

3 Web System Architecture

The web system presented in this paper is modular, client-server, with a template system for its appearance; and it is based on the JSP technology, and uses a MySQL database. JSP and MySQL were chosen mainly because they are robust, portable, and widely used in the market. The system architecture is supported on the Model-View-Controller (MVC) design pattern, and the Responsive Web Design (RWD) pattern for adjusting its web interfaces to different devices. The system architecture is composed of five components: the core, the Java beans, the model, the actions and the views.

3.1 Components of the Web System

The Model Component is in charge of the data persistence; it maintains connections to the database; it carries out transactions; and executes all the necessary methods for maintaining the data in the database. This component is divided in three subcomponents: a model class with all the methods for accessing data; a web model class, which is in charge of obtaining data from the application context; and the init model class, which initializes all the necessary information before establishing connections.

The Action Components are in charge of the communication with the View Components. They are implemented as classes with some attributes that obtain requests from clients, pass requests to the controller, obtain context from the application, and provide the model of the application. Following the MVC design pattern, the Action Components work directly with the model and return a response to the controller.

The Bean Components allow creating specific objects for manipulating data from the database. There are bean components for users, products, categories, and for any other entity in the system; which have *set* and *get* methods for reading and setting their attributes; they also have a method to retrieve their *string* representation. The bean components obtain data as follows: views request user data to the controller; the controller searches and executes the corresponding action; the action returns a result set object; a method in the bean component converts the result set into the type of the bean, and returns it to the view; the view uses this object as any other bean. It should be noticed that the data is encapsulated in the bean component.

The View Components are the interface to the user of the system. These components are implemented with XHTML and JSP. The responses from the controller are actually calls to JSP; they obtain parameters that allow them displaying information. An special function of the system presented in this paper, is the ability of obtaining information from the model without the need of an action bean.

It should be noticed that the Action and Bean Components were incorporated because they allow delegating intermediary actions between controller and model.

3.2 Model-View-Controller (MVC)

The Model-View-Controller (MVC) design pattern is an ordered set of subsystems, which are composed of other systems that are below them, and provide the base of the implementation for those systems that are above. The objects in each layer are normally independent, and there is a client-server relationship among the bottom and the top layers. The system presented in this paper is based on the MVC design pattern, which is divided in three layers: the view, the controller and the model. The *View* is the presentation layer, which provides the graphic interface and it is in charge of the presentation of the data to the user, as well as to collect information from the user. The *Controller* is the business layer, which provides the functionality of the application; it is the layer where user and system interchange messages; it is also in charge of managing the requests from the user, and executing the necessary operations to solve such requests. The *Model* is the persistent layer; it stores all the information of the system; it also allows access to the information in a secure and controlled way; this

level is composed of the database and a class called model, which contains all the methods that provide access to the data. The MVC design pattern was chosen because it allows modifying the layers independently, without affecting the rest of the system.

3.3 Responsive Web Design (RWD)

The Responsive Web Design (RWD) is a recently created pattern; it has been adopted by web developers, because it tries to solve the problems that exist when creating web pages for different devices. Figure 1 shows a web page in different resolutions, where the elements of the page are distributed according to the resolution of the device where the web page is being displayed.

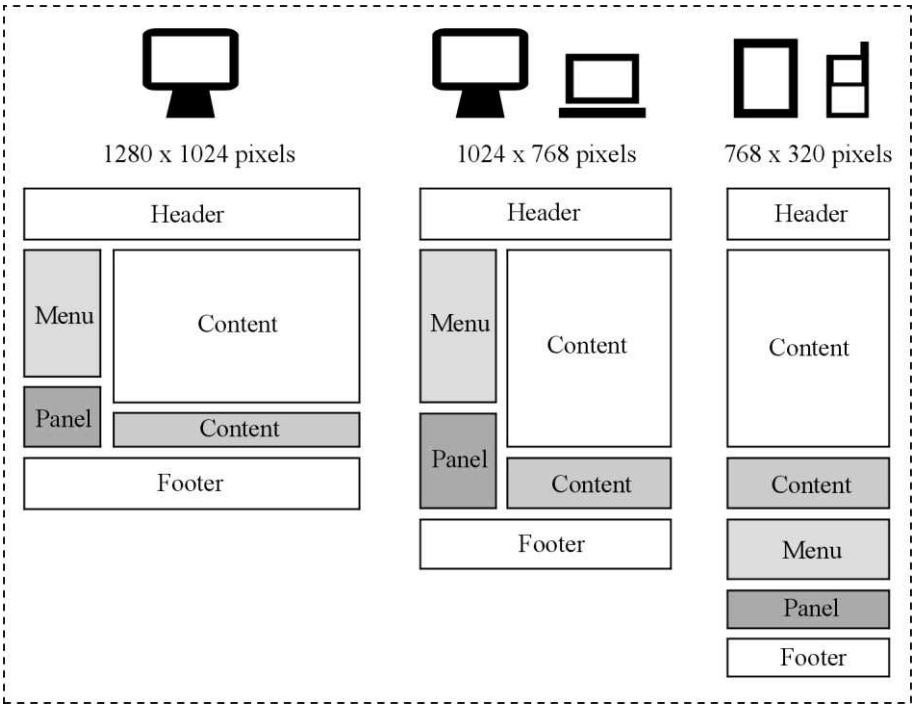


Fig. 1. Distribution of elements on the web page according to the resolution.

This RWD pattern is centered in a unique design that is adaptable to different devices and their resolutions, without developing specific implementations; it is mainly based on Media Queries (MQ), which are part of the last CSS 3 specification, and allow changing styles according to the resolution of the device where the web page is visualized. For example, if a web page is designed for a resolution of 1024 pixels, the font size would be 12 pixels; however, when changing the resolution, the font size would be the same size, making it difficult to read the web page. MQ can determine parts of the CSS code to appear, or change attributes of classes depending on the resolution. This way, there can be sections of CSS code for different resolutions.

4 Web System Implementation

This section provides details about the implementation of the system. It first provides a site map with all the web pages that compose the system; it shows the interface of some parts of the system; it also describes some tests on different devices; and finally it describes the technologies and tools used for developing the system.

4.1 Site Map

Figure 2 shows the site map of the web system, which is considered to be simple and organized, compared to those of the products analyzed. It considers three different users: administrators, registered and non-registered users of the e-commerce web site.

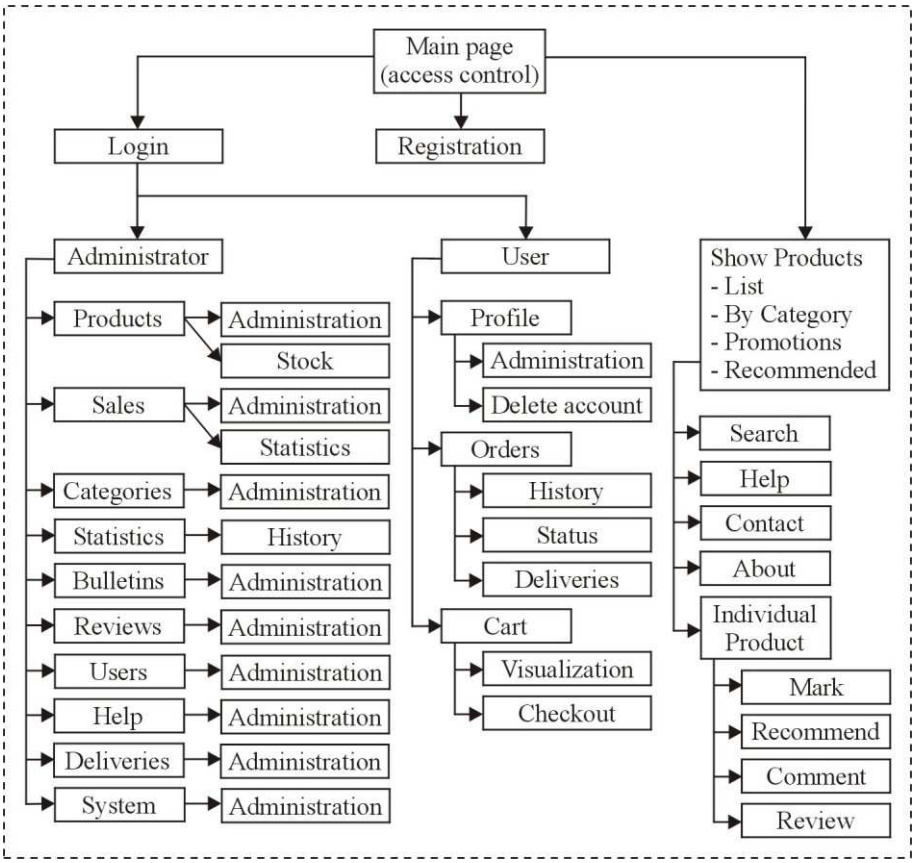


Fig. 2. Map of the web site for administrators, registered and non-registered users.

The left side of the site map illustrates the actions that an administrator can carry out on the system, such as administering products, sales, categories, statistics, bulletins, reviews, users, and deliveries; the center of the figure illustrates the actions that a

registered user can carry out on the system, such as editing its profile, modifying its orders, and its shopping cart; and the right side of the figure illustrates the actions that non-registered users, administrators, and registered users can carry out on the system, such as visualizing products, searching products, viewing promotions and recommendations, marking products, recommending products, commenting on products, reviewing products, looking for help, and contacting the web site administrator.

4.2 Web Interface

This subsection provides some screenshots of an e-commerce web site generated with the web system presented in this paper. The screenshots show a web page with a list of products, in three different resolutions, according to the device where the web site is being displayed. Figure 3 shows the web interface displayed on a desktop computer with a resolution of 1240 x 1024 pixels, which is the maximum; and Figure 4 shows two web interfaces, the first one displayed on a tablet with a resolution of 960 x 720 pixels, and the second one displayed on a smartphone with a resolution of 320 x 480 pixels. The web page is divided in two panels: the left panel shows the menus for all the categories of products, news, reviews, bulletins, among others; and the right panel visualizes the list of products, with their ids, names, category, stock, price, state, and some actions. The screenshots in Figure 4 show the same web page, but with reduced information and simplified menus, due to the resolution of the devices used. For example, in the web page shown on the smartphone, the left panel is hidden, but can be displayed through a button on the top of the page.

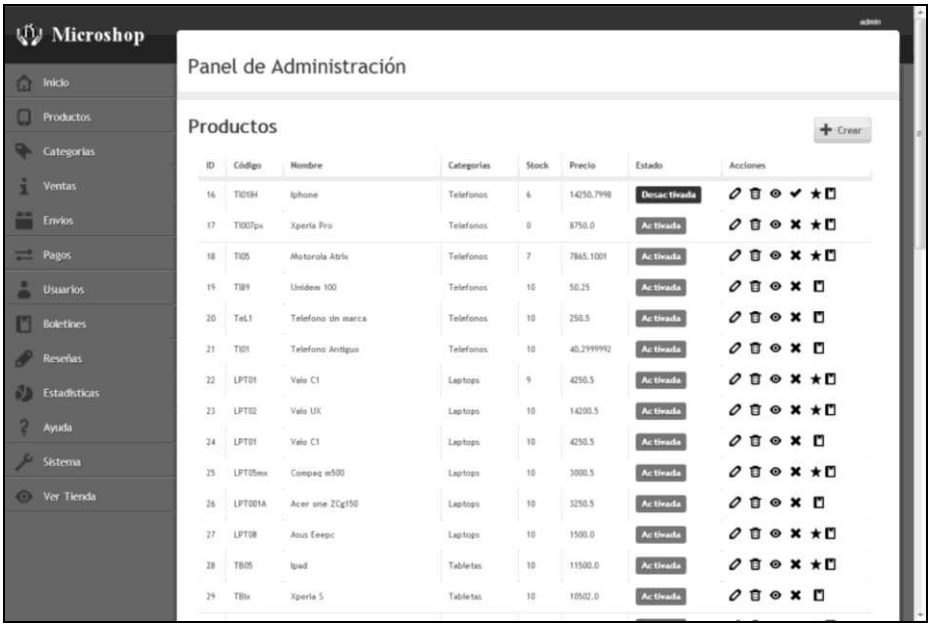


Fig. 3. E-commerce web site displayed on a desktop computer (1240x1024).

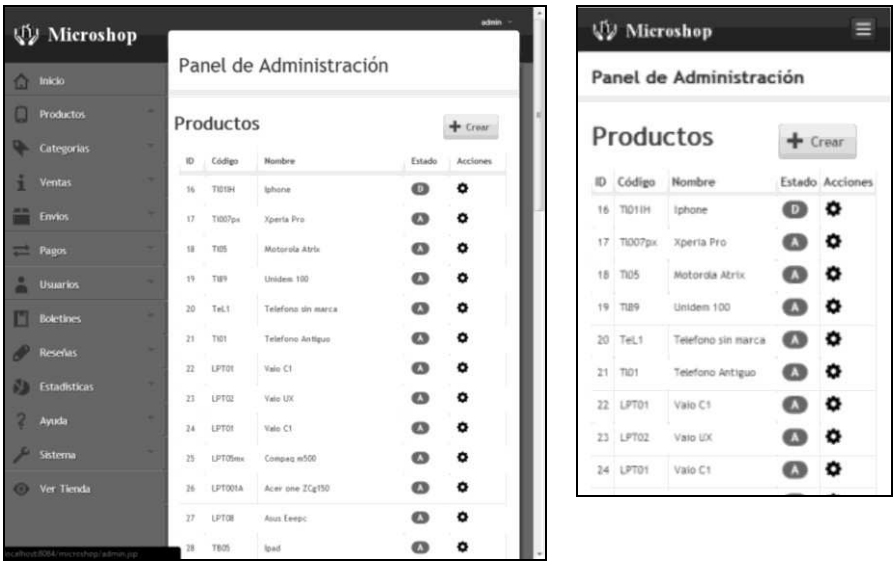


Fig. 3. E-commerce web site displayed on a tablet (960x720), and on a smartphone (320x480).

4.3 Tests on Different Devices

The functionality of the web system was tested by generating an e-commerce web site, and testing every functionality on three different devices: a desktop Intel i7, 8GB RAM, Windows 7 Operating System, with a monitor with resolution of 1024 x 1240 pixels; a tablet HP Touch Pad SnapDragon 1.2GHz, 1GB RAM, Android Ice Cream Sandwich Operating System, with a resolution of 960 x 720 pixels; and a Smarthphone Sony Xperia mini pro SnapDragon, 1.6Ghz, 1GB RAM, Android Ice Cream Sandwich Operating System, with a resolution of 320 x 480 pixels. The web browsers used were the following: Firefox V.17, Google Chrome V.25, and Internet Explorer 8, for the desktop computer; and Firefox Mobile, Google Chrome Mobile, and Android web browser, for the tablet and the smartphone.

4.4 Technologies and Tools Used

The main technologies and languages used for the implementation of the web system were the following: Apache Tomcat web server, JSP, servlets and Java beans; MySQL database management system; HTML and XHTML for the web page content; CSS for the presentation and appearance of the web pages; JavaScript for validations and dynamic generation of content for the web pages; JQuery framework for some of the functions that facilitate some tasks; Twitter Boostrap framework for the creation of complex structures, such as predefined columns and basic components; Java beans as reusable software components; servlets for the business logic of the system, and the implementation of the controller; and JSP for processing information and generating dynamically the web interface of the system.

The tools used for the implementation of the system are free and open source: Star UML was used for creating UML diagrams for the analysis and design stages; NetBeans as an IDE for creating and editing Java classes, JSP, CSS and JavaScript files; Apache Tomcat web server for deploying the web application; DBDesigner as the IDE for modeling the database, creating an entity-relation model, and a relational model; Web Developer Toolbar is a plug-in for different web browsers, which has debugging tools, and other web page design facilities, such as the function to resize the web pages to test different resolutions without closing the web browser; ResizemyBrowser is a web page that resizes the web browser to a specific resolution, which allows testing for mobile devices.

5 Conclusions and Future Work

This paper presented a configurable web system for generating e-commerce web sites, which can be displayed on mobile devices adjusting the resolution automatically. We have successfully tested the functionality of this prototype on three different devices: a desktop computer with the resolution of 1240x1024 pixels, a tablet with a resolution of 960x720 pixels, and a smartphone with a resolution of 320x480 pixels.

The architecture of the system is client-server, supported on the Model View Controller design pattern for the internal functionality, using a controller, a model, and several actions; and the Responsive Web Design pattern for the web interface. They have proved to be design patterns that adapt to these types of applications. The architecture of the system is composed of five components: the core, the Java beans, the model, the actions and the views. A site map of the web system was also presented with all the functionality for administrators, registered and non-registered users.

It is needed to carry out more functionality and usability tests with final users, and administrators of e-commerce web sites, in order to evaluate the system and adjust it accordingly. As part of the future work, it is also planned to incorporate a recommender system for products, based on user preferences and profiles.

References

1. OsCommerce. Available at: http://www.oscommerce.com/. Last accessed in June 2014.
2. Magento. Available at: http://www.magento.com/. Last accessed in June 2014.
3. Prestashop. Available at: http://www.prestashop.com/. Last accessed in June 2014.
4. Opencart. Available at: http://www.opencart.org/. Last accessed in June 2014.
5. Comparative table of shopping cart software. Available at: http://en.wikipedia.org/wiki/Comparison_of_shopping_cart_software. Last accessed in June 2014.
6. Zen-Cart. Available at: http://www.zen-cart.com/. Last accessed in June 2014.
7. Go eCommerce. Available at: http://www. go-ecommerce.com/. Last accessed in June 2014.
8. Shoppizer. Available at: http://www.shoppizer.com/. Last accessed in June 2014.
9. Konakart. Available at: http://www.konakart.com/. Last accessed in June 2014.
10. Jada-Site. Available at: http://www.jadasite.com. Last accessed in June 2014.

A visual aid on the ecological impact of products for online retail

Sandra Cecilia Urbina Coronado, Satu Elisa Schaeffer[1], and Sara Elena Garza Villarreal

Facultad de Ingeniería Mecánica y Eléctrica, Universidad Autónoma de Nuevo León, San Nicolás de los Garza, Nuevo León, Mexico,
[1] Corresponding author: elisa.schaeffer@uanl.edu.mx

Abstract. We propose a browser plug-in for online retail that visualizes information regarding the ecological impact of a product to aide the user in making "greener" purchases. We document the design and implementation of such a tool, outlining also the design for a back-end recommendation engine, and report user experiments on the proposed plug-in, with favorable results.

Keywords: Eco-labels, shopping assistant, usability, ecology, browser plug-in, e-commerce

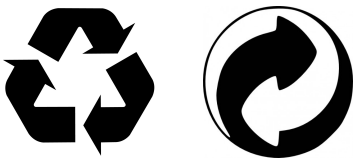
1 Introduction

The use of information technology already assists people in many of their everyday tasks, at home and at the office, helping human beings avoid errors and increase their efficiency. Shopping online is very common and more and more people prefer to shop for groceries online and have them delivered home instead of going to a supermarket themselves [9]. This opens a window of opportunity for *informing* the consumer of the products at the moment of (considering) purchase.

Interesting information could include details such as food calorie intake or the presence of allergenic agents (although the legal implications of failing such a warning make this a risky task for a service provider), and in this work we propose a prototype for providing the user of an online retail store with information regarding the *ecological impact* of the products viewed with the web browser. The goal is to guide the consumers towards environmentally friendly products and hence help them be "green" without the increased cognitive load implied in having to remember to check whether a detergent is non-toxic and if a packaging material is recyclable, for example.

The proposed browser plug-in could easily be adapted to provide dietary (calories, ingredients), financial (budget concerns, unit prices when not shown), or ethical (use of child labor) information about a product and/or its manufacturer, but we chose to begin with ecological information as the first case study due to the pressing need in developing countries to elevate the consciousness of the consumers to the importance of preserving the nature and the climate [4].

©L. Berthet, J. C. Chimal, E. A. Santos, E. Castillo
Advances in Computer Science
Research in computing Science 81, 2014 pp. 103-113



(a) Recycle. (b) Recollect.

Fig. 1: Eco-labels used to indicate characteristics of product packaging.

Whether or not a product is eco-friendly depends on its useful life (for how long will it be used before being discarded), the implied waste (packaging and remains of the product itself), the manufacturing (ingredients, energy usage, contaminants produced in the process), and also the transport of the product from the place of origin to the place of purchase (local products are likely to contaminate less during transport).

Our proposed system processes the web page displayed in a browser upon viewing a product in an online retail site, extracts information on the product to consult it with a recommendation back-end, and then visualizes the received ecological information in the browser for the user to view and consider. The rest of the paper is organized as follows: Section 2 briefly touches the background to establish terminology and context for the present work, and then Section 3 discusses related work, after which Section 4 describes the proposed solution in terms of the software involved and the functionality provided. Then in Section 5 we describe and report the experiments carried out with the implemented prototype, and finally in Section 6 we conclude the present work and discuss directions for future work.

2 Background

There are many ways to describing consumer behaviour as ecologically responsible, all implying that the purchase decisions are made considering the ecological impact of the product, its entire manufacturing process, the waste generated by the product itself, its packaging, the transport of materials, and the final product [13]. We refer to consumers with heightened awareness of these issues as *green consumers* and to the products they find appropriate as *green products* [6]. Among the main factors in determining the "greenness" of a product lies also the energy consumption and that of clean water [8].

In order to help green consumers identify green products, several organizations maintain and enforce normativity of ecological excellence of some sense, both to certify the manufacturing phase and to ensure eco-friendly characteristics of the final product. Contamination is to be avoided in obtaining and transporting the raw material, in the process of manufacturing the product, in the transport of the product to the consumer, and also in the process of consumption and the resulting waste.

The conformity to a certain ecological norm or (voluntary) regulation is often indicated by introducing an *ecological label* (eco-label for brevity) in the packaging of the product. These are commonly logotype-like figures, accompanied possibly by a brief text. A typical example is the symbol indicating that the packaging material can be recycled (c.f. Figure 1a), and also one indicating that the packaging is reusable (especially for plastic and glass bottles) is frequently used (c.f. Figure 1b). There are hundreds of eco-labels worldwide. Especially the paper industry has been active in the field; *FSC certified* by the Forest Stewardship Council, for example, indicates a proper use of the forests [11]. Another active field of eco-labels is that of the *carbon footprint*, promoted by the Carbon Trust (http://www.carbontrust.com/).

3 Related work

Published literature on consumer behavior as such is abundant in recent years. Wang [17] analyzes the relationship between consumer lifestyle and purchase decisions in order to identify green-consumer profiles among a sample of 433 persons (16 years or older), whereas Wang [16] explores the connection between lifestyle and green marketing strategies; both studies were carried out in Taiwan. Clare et al. [3] analyze factors affecting green consumption in Australia. A study of the effect of ecological beliefs on purchase decisions in Iran is presented by Hossein Khorshidi et al. [10].

The effect of consumer consciousness of ecological aspect on their purchases in practice is discussed by Pereira et al. [15], and also Juwaheer and Pudaruth [12] take a marketing approach to studying and exploiting consumption patterns to promote ecological lifestyle. The positive effect of marketing on green consumption is demonstrated by Kumar [14].

In their study, Bosch and Valor [2] found that consumers struggle to understand eco-labelling and are unable to assess the credibility of ecological information on products; only the consumers who express the most interest in ecology are unhindered in their comprehension. Also Grunert et al. [7] report a study on ecological labels, with emphasis on food products. The focus of the work proposed in the present work is making ecological information more readily available and easier to interpret for the general population of consumers, not only those that are motivated enough to inform themselves.

4 Proposed solution

The development of the software took place in four stages: planning, design, implementation, and testing [1]. The intended users were defined as adults between ages 18 and 40 with an interest in using information technology in their daily lives and who shop for groceries at least once a month. Due to the availability of an API for developing plug-ins for the Chrome browser by Google[1], we chose

[1] https://www.google.com/intl/en/chrome

to implement the prototype for that platform, and also expect the target users to be familiar with Chrome and with plug-ins in general.

The main functionality for the plug-in is the interactive visualization of ecological impact of the currently selected product. This needs to be done in a visible, but non-disruptive way that the user perceived as pleasant and informative. Technically this is achieved by *code injection*[2] to the websites of the stores to include the ecological information within the store website. Some stores block injection of code to their content, in which case we would have to access the store via a proxy, which in turn would have to be re-directed to the original site for the actual purchase. For the prototype we chose to work with sites that permitted code injection as such to avoid this issue.

Architecturally, the process is as follows:

1. Upon selecting a product, the plug-in identifies the information associated to it and creates a query from it.
2. The query is sent to a recommendation-engine back-end running on a separate web server.
3. The server processes the query to produce a response with the ecological information available, if any.
4. The plug-in adapts the styles and the sizes on the website to accommodate the response within the view of the browser.
5. The user may interact with the visualization of the plug-in to access further details.

The user is given the liberty to hide the plug-in at any time, as well as to customize the preferences regarding the type of information shown in the plug-in.

4.1 Product recognition

Upon activating a specific *division* (`<div>` in HTML5) on a web site, the plug-in processes the HTML content of that division to attempt to identify a product. The generalization of this to be automated over several websites is a machine-learning task, and we discuss here the particular case of the web store of the Soriana supermarket chain in Mexico. The divisions for showing products at the Soriana web store are all identified as `carR2`, and hence an event for the mouse-over on such a section is easily created. Within that element, the product name is placed within a `font`-directive and can be extracted with a simple HTML parser.

The design of the plug-in was done with *Bootstrap*[3] to obtain a responsive design. The back-end server is based on the *Google App Engine*[4]. The implementation with the API for Chrome plug-ins involves HTML5, JavaScript, and CSS3.

[2] A mechanism that inserts program code into a web page to modify its functionality or the way it looks.
[3] https://getbootstrap.com/
[4] https://developers.google.com/appengine/

Fig. 2: Principal elements of the plug-in.

A *content script* is used to detect when the user enters a compatible web store (such as the Soriana web site) and begins reacting to the mouse-over events while the user remains on that same site. Also the user location is retrieved[5] and the distance between the user and the place of origin of the product is computed to estimate the ecological impact of the transport involved. The plug-in is described in a file called `manifest.json`, determining its permissions and the scripts it is allowed to access whereas the libraries of Chrome are utilized in `background.js`; Figure 2 shows the principal elements of the plug-in.

On the server side in *Google App Engine*, a simple `GET` suffices to receive the query parameters. A web application written in *Python*[6] processes the resulting

[5] The geolocalization of HTML5 is used; user permission is requested explicitly.
[6] https://www.python.org/

ID	name
4644337115725824	Carranco Enchiladas Potosin...
4785074604081152	Hamburguesa Western Griller...
4925812092436480	Tyson Alitas Picantes 700 B...
5066549580791808	Soriana Hamburguesa De Res ...
5207287069147136	Older Grill Carne P/Hamburg...
5348024557502464	Soriana Pizza Indiv Queso 1...
5629499534213120	"aceite nutrioli"
5770237022568448	Older Grill Carne Para Hamb...

Fig. 3: An extract of the server-side database.

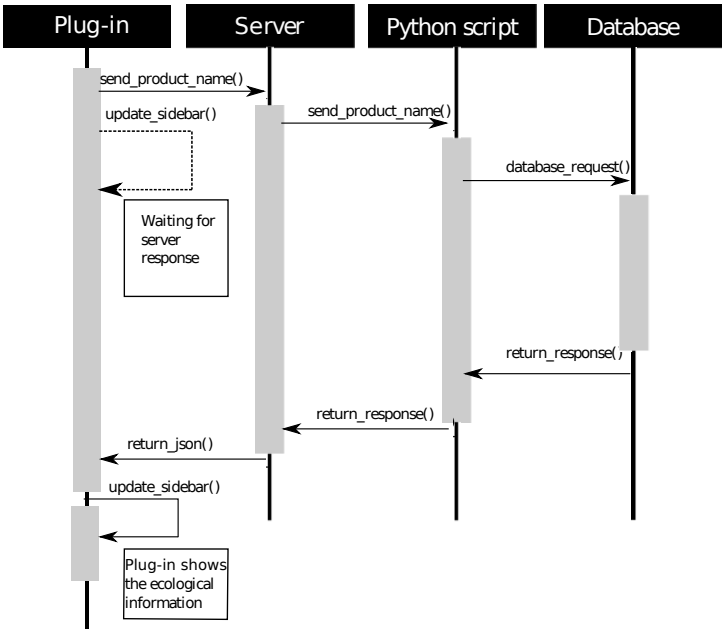


Fig. 4: Sequence diagram of the client-server communication.

HTTP request (produced in the plug-in with `AJAX`[7], checks if the product already exists in the recommendation engine (in our prototype implemented as a simple data base), inserts the product if it is not found, and uses the data stored in the database (if any) to produce a HTTP reply. Figure 3 illustrates the server-side storage, which is very simple in the present prototype, whereas Figure 4 shows the sequence with which the communication takes place. The ecological scores were artificially created on a scale to ten.

For the visual design, two alternatives were implemented: a summary view (shown in Figure 5a) and an extended view (shown in Figure 5b). The categories used for the aspects of eco-friendliness are very simple, this being a first prototype and the environmental information being locally stored. Our future work includes crawling on the web for ecological data, providing a social network for recommendation sharing, as well as information modules provided by government agencies or independent organizations.

5 Experiments

We carried out an experiment to assess whether the use of the plug-in affects the purchase decisions, i.e., whether there is some sign of an elevated ecological consciousness upon its usage. The sample consists in 24 persons between 18 and

[7] https://developer.mozilla.org/en/docs/AJAX

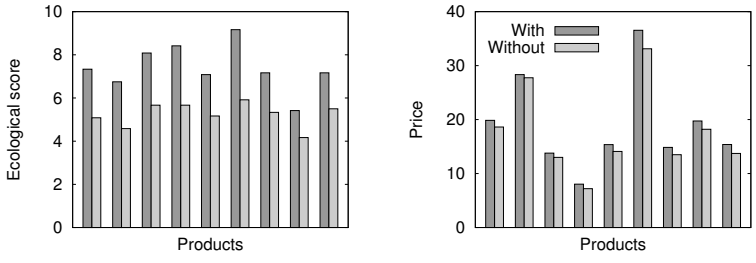
(a) Summary (pop-up). (b) Extended (sidebar).

Fig. 5: Screen shots of the plug-in used on a web store, cropped for detail.

40 years of age who have purchased something within the last month and are interested in trying out the proposed system. Each subject was provided with a list of basic products commonly acquired at a supermarket [5]. The sample was divided in two groups: one half used the Soriana website with the plug-in and the other half without it.

The task assigned to the subjects consists in purchasing products indicated on a shopping list with nine items, provided to the user. The control group uses the website only, whereas the other group employs the plug-in. The subjects are requested to select the products they would actually purchase, as honestly as possible. The plug-in stores automatically the information regarding the prices and the ecological profile of the selected products, allowing us to determine whether the users of the plug-in are in fact willing to pay more in order to select greener products. The results are shown in Figure 6.

After the website usage, those who had used the plug-in answered a questionnaire regarding their assessment of the environmental impact of their shopping habits and the usability of the plug-in on a scale from one (totally agree) to five (totally disagree). The results are shown in Table 1 and are quite favorable. The table indicates the questions posed to the users and the percentages



(a) Ecological score. (b) Price in Mexican pesos.

Fig. 6: Averages per product with and without the plug-in.

obtained for each answer option. The first set of questions were simple yes/no questions, whereas the second set was on a scale from one to five, one indicating the strongest agreement and five indicating the strongest disagreement.

Table 1: The percentage of responses per category in questionnaires.

(a) Ecological conscience.

Question	**Yes**	**No**
Do you prefer to purchase products that do not harm the environment?	75	25
When you shop, do you check if the products are reusable?	42	58
Do you feel that the plug-in helped you obtain information on the ecological impact?	83	17
Do you feel that the plug-in helped you make purchase decisions?	83	17
Do you think that the plug-in should be promoted among shoppers?	92	8

(b) Usability.

Question	**1**	**2**	**3**	**4**	**5**
The plug-in is easy to use.	75	17	0	8	0
I can easily find what I want in the plug-in.	75	17	8	0	0
I enjoy using the plug-in.	58	17	25	0	0
It is confortable to shop with the plug-in.	59	33	0	8	0
The information provided in the plug-in is relevant.	50	50	0	0	0
I would use the plug-in again in the future.	67	25	8	0	0
I would recommend the plug-in to a friend.	42	50	8	0	0
The plug-in is attractive.	59	33	0	8	0
The plug-in is simple and clear.	75	25	0	0	0

6 Conclusions

In this work we describe a prototype for providing ecological information on products while a consumer browsers an online retail store. The products are identified on the client side within a browser plug-in and then a server is consulted for information regarding the ecological impact of the product, after which the client side visualizes the received recommendation within the view of the browser. The design and implementation of the prototype is described, together with experiments that indicate a favorable effect on consumer behavior in a laboratory setting.

As future work, we leave the generalization of the present plug-in into a multi-platform tool that functions on different browsers and also on mobile devices. Also the automatization of the product recognition as well as the construction of a full-scale recommendation engine for the back-end are outside the scope of the present paper. Increasing the complexity of the system will also require more extensive usability evaluations as part of future work.

Bibliography

[1] Arrau, F., de Chile. Escuela de Ingeniería, P.U.C.: Metodología ágil y de bajo costo para la implementación del modelo CMMI en PYMES. Pontificia Universidad Católica de Chile, Chile (2008)

[2] Bosch, I.C., Valor, C.: La relación del consumidor con el etiquetado responsable. Responsabilidad Social de la Empresa 4(1), 79–104 (2012)

[3] Clare, C.D., Taghian, M., Lamb, P., Peretiatkos, R.: Green products and corporate strategy: an empirical investigation. Society and business review 1(2), 144–157 (2006)

[4] Coelho, T.M., Castro, R., Gobbo Jr., J.A.: PET containers in Brazil: Opportunities and challenges of a logistics model for post-consumer waste recycling. Resources, Conservation and Recycling 55(3), 291–299 (2011)

[5] Comité fronterizo de obreros: Lista de productos básicos (Ciudad Acuña). http://www.cfomaquiladoras.org/listaproductosbasicos.es.html (Nov 2013), [Date of access: 2014-26-06]

[6] Crea Medioambiente: El consumidor verde y los certificados ecológicos. http://www.creamedioambiente.com/blog/item/78-el-consumidor-verde-y-los-certificados-ecol%C3%B3gicos (Mar 2013), [Date of access: 2014-02-03]

[7] Grunert, K., Hieke, S., Wills, J.: Sustainability labels on food products: Consumer motivation, understanding and use. Food Policy 44(1), 177–189 (2014)

[8] Holzer, D., Media, D.: What Does Eco-Friendly Mean? (Home Guides by Demand Media). http://homeguides.sfgate.com/ecofriendly-mean-78718.html (2013), [Date of access: 2014-02-03]

[9] Horrigan, J.: Online Shopping (PewResearch Internet Project). http://www.pewinternet.org/2008/02/13/online-shopping (Feb 2008), [Date of access: 2014-23-06]

[10] Hossein Khorshidi, G., Gholizadeh, T., Naghash, A.: Consumer behavior in relation to the decision to buy green products: an investigation on Iranian consumers. International journal of management perspective 1(3), 61–71 (2013)

[11] Jeffries, A.: Is it green?: Eco-Labels and Certifications (inhabitat — design will save the world). http://inhabitat.com/is-it-green-eco-labels (Apr 2008), [Date of access: 2014-02-03]

[12] Juwaheer, T.D., Pudaruth, S.: Analysing the impact of green marketing strategies on consumer purchasing patterns in mauritius. World Journal of Entrepreneurship, Management and Sustainable Development 8(1), 36–59 (2001)

[13] Klöpffer, W.: Life cycle assessment. Environmental Science and Pollution Research 4(4), 223–228 (1997)

[14] Kumar, D.: Does green marketing works with consumers. Wulfenia journal 20(3), 329–347 (2013)

[15] Pereira, E., Mykletun, R., Hippolyte, C.: Sustainability, daily practices and vacation purchasing: are they related? Tourism Review 67(4), 40–54 (2012)
[16] Wang, W.L.: Most feasible strategies for green marketing mix under business sustainable development. The Business Review 20(1), 297–303 (2012)
[17] Wang, W.L.: A study on consumer behavior for green products from a lifestyle perspective. The Journal of American Academy of Business 18(1), 164–170 (2012)

Diseño e Implementación de Composiciones Paralelas de Alto Nivel como un Modelo de Objetos Paralelos

Mario Rossainz-López[1], Ivo H. Pineda-Torres[1], Patricia Domínguez Mirón[1], Manuel Capel Tuñón[2]

[1]Benemérita Universidad Autónoma de Puebla
Avda. San Claudio y 14 Sur, San Manuel. 72000, Puebla, Pue. México
{rossainz,ipineda}@cs.buap.mx, paty.dguez.m@gmail.com
[2]Universidad de granada, DLSI-ETSII
C/Periodista Daniel Saucedo Aranda S/N, 18071, Granada, España
manuelcapel@ugr.es

Abstract. This paper presents the design of a model for programming parallel applications in a structured way. To achieve this, the model of High Level Parallel Compositions or CPANS under the paradigm of object orientation is used to solve problems which are parallelizable algorithms, using the concept of structured parallelism and possible integration within the paradigm of orientation objects to make way for the definition and model CPAN, using the Parallel objects (PO). The types of communication of PO, the templates that define the components of a CPAN as well as the syntactic and semantic definition of a CPAN are also described in this paper.

Keywords: CPAN, Programación Paralela Estructurada, Objetos Paralelos, Programación Orientada a Objetos, Paralelismo Estructurado.

1 Introducción

Un enfoque que intenta atacar el problema de la paralelización de algoritmos y programas consiste en tratar de hacer que el usuario desarrolle sus programas según un estilo de programación secuencial y ayudado de un sistema o entorno específico de programación, intentar producir su contraparte paralela automáticamente [1], [6]. Existen, sin embargo, dificultades de implementación intrínsecas a la definición de la semántica formal de los lenguajes de programación que impiden la paralelización automática sin contar con la participación del usuario, por lo que el problema de generar paralelismo de manera automática para una aplicación de propósito general sigue considerándose un problema abierto. Un enfoque alternativo es el paralelismo estructurado. Las aplicaciones paralelas generalmente siguen patrones predeterminados de ejecución, que son rara vez arbitrarios y no estructurados en su lógica [4]. Las Composiciones Paralelas de Alto Nivel o CPANS son patrones paralelos de comunicación bien definidos y lógicamente estructurados que, una vez identificados en términos de sus componentes y de su esquema de comunicación, pueden llevarse a la práctica como constructos añadidos a un lenguaje de programación orientado a objetos y estar

©L. Berthet, J. C. Chimal, E. A. Santos, E. Castillo
Advances in Computer Science
Research in computing Science 81, 2014 pp. 113-123

disponibles como abstracciones de alto nivel en las aplicaciones del usuario. Esto es posible ya que las estructuras de interconexión de procesadores más comunes, como los pipelines, los farms y los trees, pueden ser abstraídas utilizando CPANS dentro de un modelo general de desarrollo de Objetos Paralelos [3].

2 El Paralelismo Estructurado

El enfoque estructurado para la programación paralela se basa en el uso de patrones de comunicación/interacción (pipelines, farms, trees, etc.) predefinidos entre los procesos de una aplicación de usuario. Éste permite diseñar aplicaciones en términos de CPANS capaces de implementar los patrones ya mencionados. La encapsulación de un CPAN debe seguir el principio de modularidad y debe proporcionar una base para obtener la reusabillidad del comportamiento paralelo de la entidad software que éste implementa. Se crea entonces un patrón paralelo genérico que proporciona una posible representación de la interacción entre los procesos independientes de la funcionalidad de estos. Así, en lugar de programar una aplicación paralela desde el principio y de programar con un nivel de detalle cercano a la capa protocolo de red, el usuario simplemente identifica los CPANS que implementan los patrones adecuados para las necesidades de comunicación de su aplicación y los utiliza junto con el código secuencial que implementa los cómputos que realizan sus procesos [4], [5].

2.1 El Paradigma de la Orientación a Objetos (OO)

El paradigma de la OO se utiliza en el presente trabajo para encapsular y abstraer patrones comunes de interacciones paralelas hacia un estilo de paralelismo estructurado. La idea es considerar a los CPANS como objetos encargados de controlar y coordinar la ejecución de sus componentes internos proporcionando características importantes de la OO tales como la uniformidad, la genericidad y la reusabilidad.

3 La definición del modelo CPAN

El objetivo es la de representar cualquier tipo de patrones paralelos de comunicación entre los procesos de un algoritmo paralelo y distribuido como clases, siguiendo el paradigma de la Orientación a Objetos [2]. Un CPAN es una composición de un conjunto de objetos de tres tipos. A partir de dichas clases, un objeto puede ser instanciado y la ejecución de un método del objeto se puede llevar a cabo a través de una petición de servicio (ver fig. 1). Los tipos de objetos que componen un CPAN son:
- **un objeto manager** que controla las referencias de un conjunto de objetos Collector y Stages, que representan los componentes del CPAN y cuya ejecución se lleva a cabo en paralelo y debe ser coordinada por el propio manager.
- **los objetos Stage** que son objetos encargados de encapsular una interfaz tipo cliente-servidor que se establece entre el manager y los objetos esclavos (entidades externas que contienen el algoritmo secuencial que constituye la solución del problema), y de proporcionar la conexión necesaria entre ellos para implementar la semántica del patrón de comunicación que se pretende definir.

- **y un objeto Collector** que es un objeto encargado de almacenar en paralelo los resultados que le lleguen de los objetos stage que tenga conectados.

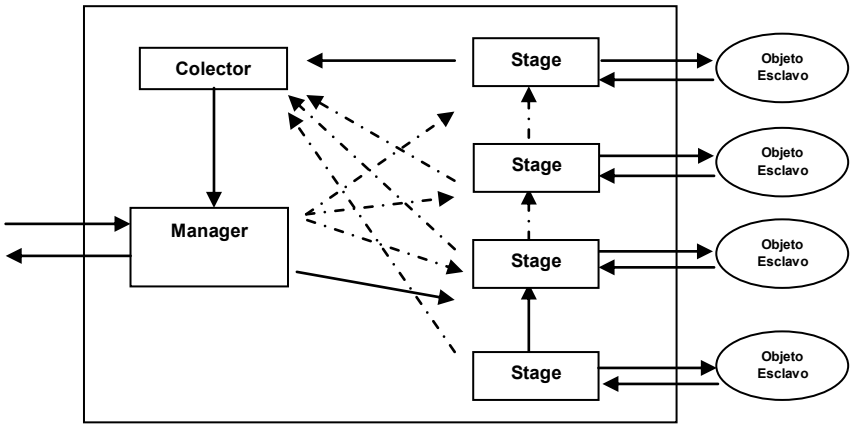


Fig. 1. Estructura Interna de un CPAN (Composición de sus Componentes)

3.1 El CPAN visto como una composición de objetos paralelos

Los objetos manager, collector y stages se engloban dentro de la definición de Objeto Paralelo (PO) [3], [4]. Los Objetos Paralelos son objetos activos, con capacidad de ejecución en sí mismos. Las aplicaciones dentro del modelo PO pueden explotar tanto el paralelismo entre objetos, como el paralelismo interno de ellos. Un objeto PO tiene una estructura similar a la de un objeto en Smalltalk, pero además incluye una política de planificación determinada a priori que especifica la forma de sincronizar una o más operaciones de la clase del objeto susceptibles de invocarse en paralelo (restricciones de sincronización, exclusión mutua y paralelismo máximo).

3.1.1 Definición Sintáctica

Desde el punto de vista sintáctico, la definición de la clase que define un Objeto Paralelo sigue el siguiente patrón o template:

```
CLASS: <nombre de la clase> INHERITS_FROM: <lista de los
nombres de clases padres directas>
    INSTANCE_STATE: <lista de los nombres y tipos de las
                     variables de instancia>
    INSTANCE_METHODS: <lista de métodos públicos>
    PRIVATE_METHODS: <lista de métodos privados>
    SCHEDULING_PART: <lista de restricciones de
                      sincronización>
END_CLASS <nombre de la clase>
```

3.1.2 Tipos de Comunicación en los Objetos Paralelos

1. El modo síncrono detiene la actividad del cliente hasta que el objeto activo servidor le cede la respuesta.
2. El modo asíncrono no fuerza la espera en la actividad del cliente. El cliente envía simplemente la petición al objeto servidor activo y continúa su ejecución. Su uso en la programación de aplicaciones también resulta fácil. Sólo hay que crear un hilo y lanzarlo para su ejecución.
3. El modo futuro asíncrono hace esperar la actividad del cliente sólo cuando, dentro de su código, se necesita el resultado del método para evaluar una expresión, [7], [4], [8]. Su uso también es sencillo, aunque su implementación requiere de un cuidado especial para conseguir un constructo con la semántica deseada.

4 Definición Sintáctica y Semántica de las clases base de un CPAN

En los PO las clases básicas necesarias para definir los objetos manager, colector, stages de un CPAN son: la clase abstracta ComponentManager, la clase abstracta ComponentStage y la clase concreta ComponentCollector. Las definiciones sintácticas siguientes han sido escritas utilizando una gramática libre de contexto publicada en [9].

4.1 La clase abstracta ComponentManager

Define la estructura genérica del componente manager de un CPAN, de donde se derivarán todos los objetos manager concretos, dependiendo del comportamiento paralelo que se contemple para la creación de cada CPAN considerado en la aplicación.

```
CLASS ABSTRACT ComponentManager
  {
    ComponentStage[] stages;

    PUBLIC VOID init (ASOCIACION[ ] list) {ABSTRACT;}

    PUBLIC ANYTYPE execution(ANYTYPE datain)
      VAR
         ComponentCollector res;
       {
        res = ComponentCollector CREATE();
        commandStages(datain,res);
        RETURN res.get();
       }

    PRIVATE VOID commandStages(ANYTYPE datain,
                               ComponentCollector res)
```

```
        {ABSTRACT;}

    MAXPAR (execution);
  };
```

4.2 La clase abstracta ComponentStage

Define la estructura genérica del componente stage de un CPAN, así como de sus interconexiones, de donde se derivarán todos los stages concretos dependiendo del comportamiento paralelo que se contemple en la creación del CPAN.

```
CLASS ABSTRACT  ComponentStage
  {
   ComponentStage[] otherstages;
   BOOL am_i_last;
   METHOD meth;
   OBJECT obj;

   PUBLIC VOID init (ASOCIACION[ ] list)
      VAR
       ASOCIACION item;
     {
      item = HEAD(list);
      obj = item.obj;
      meth= item.meth;
      if (TAIL(list) == NULL)
      am_i_last = true;
     }

   PUBLIC VOID request (ANYTYPE datain,
                                ComponentCollector res)
      VAR
       ANYTYPE dataout;
     {
      dataout = EVAL (obj,meth, datain);
      IF (am_i_last)
         TREAD res.put(dataout)
      ELSE commandOtherStages (dataout, res);
     }

   PRIVATE VOID commandOtherStages(ANYTYPE dataout,
                                ComponentCollector res)
      {ABSTRACT;}
   MAXPAR (request);
  };
```

4.3 La clase concreta ComponentCollector

Define la estructura concreta del componente collector de cualquier CPAN. Este componente implementa básicamente un buffer multi-item donde se irán almacenando los resultados de los stages que hagan referencia al collector. De esta forma se puede obtener el resultado del cálculo iniciado por el manager.

```
CLASS CONCRETE ComponentCollector
  {
   VAR
    ANYTYPE[] content;

   PUBLIC VOID put (ANYTYPE item)
     {  CONS(content, item); }

   PUBLIC ANYTYPE get()
      VAR
       ANYTYPE result;
     {
      result = HEAD(content[]);
      content = TAIL(content[]);
      RETURN result;
     }
   SYNC(put,get);
   MUTEX(put);
   MUTEX(get);
  };
```

4.4 Las restricciones de sincronización MaxPar, Mutex y Sync

Cuando se producen peticiones paralelas de servicio en un CPAN, resulta necesario disponer de mecanismos de sincronización para que los objetos que lo constituyen puedan gestionar varios flujos de ejecución concurrentemente y se garantice al mismo tiempo la consistencia de los datos que se están procesando. Para conseguirlo, dentro de cualquier CPAN, se pueden utilizar las restricciones MAXPAR, MUTEX y SYNC, para la correcta programación de sus métodos.

4.4.1 MaxPar

El paralelismo máximo, o MaxPar, es una variable de carácter permanente que indica el número máximo de procesos que pueden ejecutarse al mismo tiempo dentro de un componente en el modelo que estamos utilizando. Dicho de otra forma, el MAXPAR referido a una función del componente representa el número máximo de procesos que pueden ejecutar esa función concurrentemente. El algoritmo que ha servido para llevar a cabo correctamente la restricción de sincronización MAXPAR es el siguiente:

1. Incorporar una variable de clase en la clase de interés e inicializarla a cero.
2. Cada vez que se genere una instancia de esa clase, hay que actualizar su variable de clase sumándole uno al valor que tenía anteriormente, de forma que el número de instancias creadas corresponderá al número máximo de procesos que puedan ejecutar una función particular
3. En la función de interés implementar un semáforo de la siguiente forma:
   - 3.1. Si n es el número máximo de procesos que pueden ejecutar la función y si k es el número de procesos que actualmente se están ejecutando entonces:
   - 3.2. Si k≤n ejecutar en paralelo los k procesos, sino bloquear los k-n procesos y no despertarlos hasta que por lo menos uno de los n procesos que se están ejecutando haya terminado de forma que siempre se mantenga el límite de n procesos ejecutándose concurrentemente.

4.4.2 Mutex

La restricción de sincronización mutex lleva a cabo una exclusión mutua entre procesos que quieren acceder a un objeto compartido. Un mutex es una entidad de sincronización de Programación Concurrente que preserva secciones críticas de código para ser ejecutadas por un solo proceso cada vez, así como permitirle obtener acceso exclusivo a los recursos. El algoritmo que implementa la restricción de sincronización Mutex es el siguiente:

1. Utilizar una variable de condición que funcione como una bandera y un candado mutex.
2. Inicializar la bandera a falso
3. En la función particular de interés implementar un semáforo de la siguiente forma:
   - 3.1. Aquel proceso que adquiera el candado pondrá la bandera a verdadero y podrá ejecutar la función de interés
   - 3.2. Mientras la bandera sea cierta todos los demás procesos que intenten adquirir el candado para ejecutar la función se bloquearán y esperarán a que la bandera sea falsa para que alguno de ellos pueda tomar el candado
   - 3.3. El proceso que actualmente posee el candado podrá ejecutarse
   - 3.4. Al terminar su ejecución, pondrá la bandera a falso y liberará el candado

3.5.Aquel proceso que este al principio en la cola de espera, adquirirá el candado repitiéndose la secuencia desde el paso 3.1.

4.4.3 Sync

La restricción SYNC no es más que una sincronización del tipo productor/consumidor . El algoritmo de implementación es el siguiente:

1.Utilizar una variable de condición (digamos, n_items) que contabilice el número de elementos colocados en un contenedor.
2.Inicializar n_items a cero
3.Implementar un semáforo compartido por las funciones productor y consumidor de la siguiente forma:
  3.1. Un proceso consumidor permanecerá bloqueado mientras n_items sea igual a cero, pues con ello se indica que el proceso productor no ha colocado ningún dato en el contenedor que comparten.
  3.2. Un proceso productor colocará un dato en el contendor, adquirirá el candado y aumentará en uno el valor de n_items.
  3.3. Despertará al proceso consumidor y continuará ejecutándose en caso de procesar mas datos
  3.4. En el momento en que el n_items sea mayor que cero, el consumidor adquirirá el candado, disminuirá en uno el valor de n_items, indicando con ello que se ha procesado un dato del contenedor, y sacará el dato.
  3.5. Mientras el n_items sea mayor a cero el consumidor se estará ejecutando en paralelo con el productor.

5 Características importantes del modelo CPAN

- Capacidad de solventar los requerimientos de servicios de los objetos en los modos de comunicación síncrono, asíncrono y futuro asíncrono;
- Los objetos que se han definido poseen paralelismo interno, es decir, capacidad de gestionar varias peticiones de forma concurrente;
- Implementación de mecanismos de sincronización con semántica diferente para atender peticiones paralelas de servicio;
- Mecanismo flexible de control de tipos, es decir, es posible asociar tipos dinámicamente a los parámetros de los métodos de los objetos.

Con este modelo se han implementado los patrones de comunicación de procesos más comúnmente utilizados: cauces o pipelines, granjas o farms y árboles o trees como CPANs para resolver problemas de ordenación, búsqueda y optimización de datos

sobre un modelo de programación de memoria compartida [10]. Además, se han reutilizado algunos de estos CPANs para proponer soluciones paralelas bajo el esquema de éste modelo y resolver problemas NP-Completos como la propuesta de una solución al problema del Agente Viajero usando la técnica de Ramificación y Poda utilizando los CPANs [11]. Finalmente se han obtenido resultados de análisis de rendimiento y aceleración (speedup) de las implementaciones citadas (ver fig. 2 y fig. 3), en una máquina paralela con arquitectura de memoria compartida-distribuida para demostrar el buen rendimiento de dichas aplicaciones, ver detalles en [12], [13], [14].

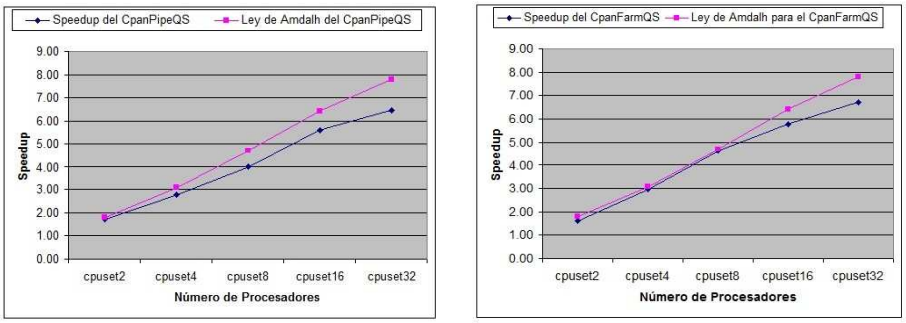


Fig. 2. Escalabilidad del Speedup para el Pipeline y Farm como CPANS con 2, 4, 8, 16 y 32 procesadores para el problema de ordenación de 50000 datos enteros

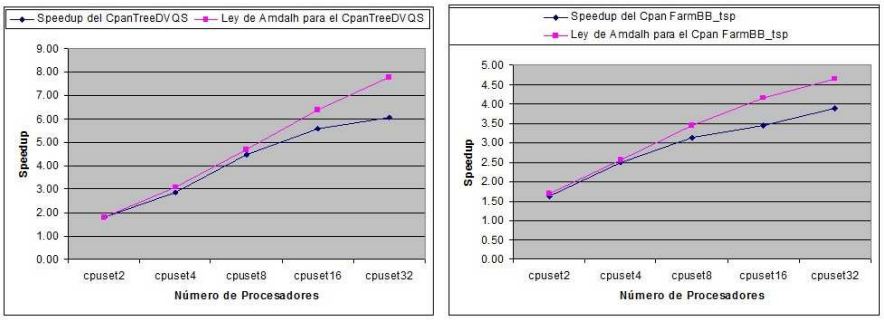


Fig. 3. Escalabilidad del Speedup para el Tree y la técnica de Branch & Bound como CPAN con 2, 4, 8, 16 y 32 procesadores para el problema de ordenación de 50000 datos enteros y del Agente Viajero para 100 ciudades respectivamente

6 Conclusiones

Se ha definido el modelo de las Composiciones Paralelas de Alto Nivel o CPANS. El modelo de los CPANS puede ser explotado para definir nuevos patrones en base a los ya construidos. Se han programado las restricciones de sincronización sugeridas por el modelo del CPAN para su funcionamiento paralelo y concurrente: MaxPar, Mutex y Sync. La programación del modo de comunicación futuro asíncrono para resultados

"futuros" dentro de los CPANS se ha llevado a cabo de manera original mediante objetos y clases. Se han programado los otros dos modos de comunicación clásicos entre procesos: el modo de comunicación síncrono y asíncrono. La programabilidad del modelo está garantizada al adoptar el enfoque basado en patrones como clases y objetos paralelos en la creación de los CPANS. La Portabilidad se da por la transparencia en la distribución de aplicaciones paralelas, lo que permite que los CPANS puedan ser portables desde algún sistema centralizado hacia un sistema distribuido sin necesidad de afectar el código fuente.

Referencias

1. Bacci, Danelutto, Pelagatti, Vaneschi: SklE: A Heterogeneous Environment for HPC Applications. Pp. 1827-52. Parallel Computing 25. (1999).
2. Brinch Hansen: Model Programs for Computational Science: A programming methodology for multicomputers. Concurrency: Practice and Experience, Vol. 5, No. 5, Pp. 407-423. (1993).
3. Corradi A., Leonardi L.: PO Constraints as tools to synchronize active objects. Pp: 42-53. Journal Object Oriented Programming 10. (1991).
4. Corradi A, Leonardo L, Zambonelli F.: Experiences toward an Object-Oriented Approach to Structured Parallel Programming. DEIS technical report no. DEIS-LIA-95-007. (1995).
5. Danelutto, M.; Orlando, S; et al.: Parallel Programming Models Based on Restricted Computation Structure Approach. Technical Report-Dpt. Informatica. Universitá de Pisa (1999).
6. Darlington et al.: Parallel Programming Using Skeleton Functions. Proceedings PARLE'93, Munich (1993).
7. Rabhi; Fethi.: A Parallel Programming Methodology based on Paradigms. In Transputer and Occam Development (18th WoTUG Technical Meeting). P. Nixon IOS Press, 239-249 (1995).
8. Roosta, Séller: Parallel Processing and Parallel Algorithms. Theory and Computation. Springer (1999).
9. Rossainz L.M.: Una Metodología de Programación Basada en Composiciones Paralelas de Alto Nivel (CPANs). Thesis (PhD). Universidad de Granada (2005).
10. Rossainz L. M, Capel T. M.: High Level Parallel Compositions (CPANs) for the Parallel Programming based on the use of Communication Patterns: 6th International Congress on ComputerScience - IPN: México (2006)
11. Rossainz L. M., Capel T. M.: Using the CPAN Branch & Bound for the Solution of Travelling Salesman Problem. Advances in Computer Science and Engineering. Volume 45. ISSN: 1870-4069. México (2010).
12. Rossainz L. M., Capel T. M.: Design and Implementation of the Branch & Bound Algorithmic Design Technique as an High Level Parallel Composition. Proceedings of International Mediterranean Modelling Multiconference. Barcelona, Spain (2006).
13. Rossainz L. M., Capel T. M.: A Parallel Programming Methodology using Communication Patterns named CPANS or Composition of Parallel Object. Proceedings of 20TH European Modeling & Simulation Symposium. Campora S. Giovanni. Italy (2008).
14. Rossainz L. M., Capel T. M.: Compositions of Parallel Object to Implement Communication Patterns. Proceedings of XXIII Jornadas de Paralelismo, pp.8-13. September 19-21. Elche, Spain (2012).

Towards a Faceted Taxonomy of Denial-of-Service Vulnerabilities in the Linux Kernel

Rolando Sánchez-Fraga, Eleazar Aguirre Anaya, Raúl Acosta Bermejo, and Moisés Salinas Rosales

Computer Research Center, National Polytechnic Institute, Federal District, Mexico
rsanchez_b12@sagitario.cic.ipn.mx, eaguirre@cic.ipn.mx, racosta@cic.ipn.mx, msalinasr@cic.ipn.mx

Abstract. The DoS vulnerabilities have become the most common type of vulnerabilities in the Linux kernel. In the last years, researches have focused on the design and implementation of detection and prevention tools of these vulnerabilities; however, despite all efforts they still prevail. This fact reveals the lack of a complete understanding of the causes and characteristics of such vulnerabilities. Because of that, we propose a faceted taxonomy, designed through the analysis and characterization of known DoS vulnerabilities to help developers and researchers organize and understand the records of actual vulnerabilities and then, focus their efforts to remove and prevent the introduction of new vulnerabilities through the development of applicable solutions.

Keywords: Taxonomy, Vulnerability, Denial of Service, Linux Kernel

1 Introduction

Despite significant advances in system protection and exploit mitigation technologies, the exploitation of software vulnerabilities persist as one of the most common methods for system compromise. Vulnerabilities are continuously discovered even in the latest versions of widely used software.

Linux kernel is a commonly attacked target. Its size and complexity are the main reasons for the existence of vulnerabilities in it. Only in 2013, almost 200 vulnerabilities in the kernel were discovered and added to the Common Vulnerabilities and Exposures (CVE) database. This is a serious problem taking in count that the OS kernel is part of the trusted computing base (TCB) and thus, any vulnerability in it could compromise the security properties of the entire system.

In the last years, the number of new CVE entries related to denial-of-service (DoS) vulnerabilities in the Linux kernel has quickly increased. From 2005 to date, each year more than half of the new CVE entries related to the Linux kernel have been DoS vulnerabilities [1].

This situation has given rise to several researches on detection of vulnerabilities via code analysis and defenses against vulnerability exploitation. However, a

©L. Berthet, J. C. Chimal, E. A. Santos, E. Castillo
Advances in Computer Science
Research in computing Science 81, 2014 pp. 123-133

complete understanding of these vulnerabilities is necessary for the development of applicable solutions.

The first step in understanding vulnerabilities is to classify them into a taxonomy based on their characteristics. A taxonomy classifies the large number of vulnerabilities into a few well defined and easily understood categories. Such classification can serve as a guiding framework for performing a systematic security assessment of a system. Suitable taxonomies play an important role in research and management because the classification of objects helps researchers understand and analyze complex domains.

Because of that, our research focuses on the design of a taxonomy for the classification of denial of service vulnerabilities in the source code of the Linux kernel through the analysis of vulnerabilities and identification of characteristics to make easier the development of applicable solution for their mitigation and identification. The main contributions of our work are the definition of the taxonomy and the results of the analysis in which it is based.

The paper is organized as follows: we examine previous work on vulnerability taxonomies in section 2, followed by a description of the methodology used in order to develop the taxonomy in section 3. Afterwards, in section 4, we present the current results of our research; the defined taxonomy, the classification of the vulnerabilities we analyze, as well as statistics related to the occurrence of certain vulnerabilities in our vulnerability set. Finally, in section 5, we conclude.

2 Related Work

Substantial research has focused on creating security taxonomies. Many of them center on different aspects of a security incident: some classify vulnerabilities, some methods of attack, and others security exploits. All previous taxonomies have one main objective; to minimize exploitable software vulnerabilities.

[2,3,4,5,6,7,8,9,10,11] presented general vulnerability taxonomies. Other works, like [12,13,14] focus on overflows and constructed C vulnerabilities taxonomies focusing on C overflow vulnerabilities. [15] and [16] center on DoS and DDoS attacks and defenses taxonomies. Most of these taxonomies were later subsequently reviewed and analyzed further, as was done by [15] and [17].

3 Methodology

Based in [18], we define in Figure 1 our methodology for the design of the taxonomy. We begin examining a subset of the vulnerabilities we want to classify. Next, we identify general characteristics of these objects. Identification of these characteristics leads to the first effort at a taxonomy. The characteristics are grouped into dimensions that form the initial taxonomy. Each dimension contains characteristics that are mutually exclusive and collectively exhaustive. This process is based on the empirical data that has been gathered about the vulnerabilities and deductive conceptualization.

Once the taxonomy is defined, we review it to look for additional conceptualizations that might not have been identified or even present in the collected data. In the process, new characteristics may be deduced that fit into existing dimensions or new dimensions may be conceptualized each with their own set of characteristics. It may even be the case that some dimension or characteristics are combined or divided. After that, we examine the vulnerabilities using the new characteristics and dimensions to determine their usefulness in classifying vulnerabilities. Out of this step comes a revised taxonomy. Then we repeat this approach, as appropriate, until we are sufficiently satisfied that the taxonomy is mutually exclusive, exhaustive, unambiguous, repeatable, accepted and useful. However, such closure is subjective and difficult to define. After the taxonomy is completed, we proceed to evaluate and test the taxonomy identifying missing or unclassified vulnerabilities.

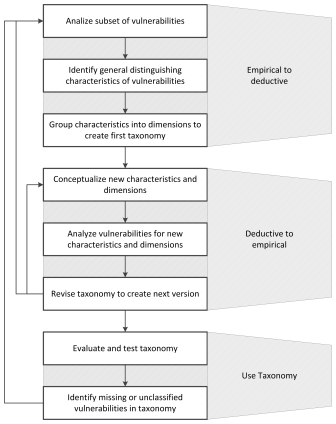


Fig. 1. Investigation Methodology

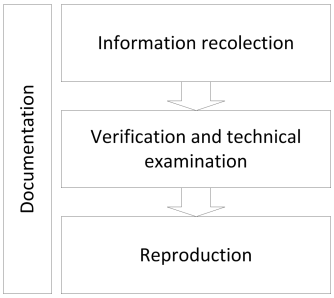


Fig. 2. Vulnerability analysis process

3.1 Analysis of the vulnerabilities

The analysis process consist of four steps [19]: information recollection, verification, reproduction and documentation as presented in Figure 2. For the recollection phase, multiple reliable sources of information of the vulnerabilities should be consulted.

The second step involves the verification and technical examination of each of the vulnerabilities. Based on the recollected information, a manual review of the source code of the kernel is performed to confirm their existence in code and determine their cause and location. Then, in the reproduction step, the objective is to create or test a proof of concept or exploit to reproduce the vulnerability and recollect additional information. Finally, all the findings must be documented. This phase is done in parallel with the other phases and ends with a *finding summary* as showed in Table 1.

Table 1. Finding Summary

Property	Value
ID	Different identifiers founded. For example, CVE and bugtraq IDs.
Localization	Localization of the vulnerability. File(s), function(s), line(s).
Description	Description of the vulnerability.
Cause	Cause of the vulnerability.
Prerequisites	Prerequisites for the vulnerability to be exploited.
Remediation	Founded remediations. Patches, modules, tools, etc.
Exploit	Existence of exploits.
CVSS Severity	CVSS Severity score.
DREAD Risk	Score according to DREAD metrics.

3.2 Evaluation and Testing

The evaluation of the taxonomy consists in proving that the taxonomy complies with the characteristics of a *well-defined* taxonomy. According to [7], satisfactory taxonomies have classification categories with the following characteristics:

- Mutually exclusive. Classifying in one category excludes all others because categories do not overlap.
- Exhaustive. Taken together, the categories include all possibilities.
- Unambiguous. Clear and precise so that classification is not uncertain, regardless of who is classifying.
- Repeatable. Repeated applications result in the same classification, regardless of who is classifying.
- Accepted. Logical and intuitive so that categories could become generally approved.
- Useful. Could be used to gain insight into the field of inquiry.

A taxonomy, however, is an approximation of reality and as such, a satisfactory taxonomy should be expected to fall short in some characteristics. Moreover, this evaluation is subjective and difficult to define. Because of that, the phase is complemented with a test, which should show the deficiencies of the taxonomy in a clearer way.

The test is formed by two parts: The classification of the set of vulnerabilities analyzed, and the classification of a set of vulnerabilities that were not analyzed. The results of these two tests will show the correctness and robustness of the taxonomy.

Finally, according with the results, missing and unclassified vulnerabilities have to be identified and evaluate if the process should be done again or the taxonomy is finished.

4 Results

In this section, we present the current results of our research. At the time of writing this paper, two iterations of the methodology shown in Figure 1 (step 1 to 8) has been performed.

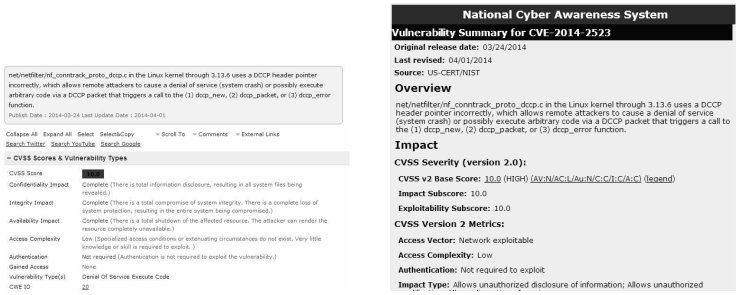
4.1 Analysis of the vulnerabilities

For the analysis, a set of 36 CVE entries of DoS vulnerabilities in the Linux kernel 3.12 were selected. These vulnerabilities comprehend the entries added in the first 8 months (between November 2013 and June 2014) of the 3.12 kernel lifetime. Linux kernel 3.12 was selected, as the latest longterm kernel at the time this research started.

For the recollection phase, multiple reliable sources were selected. We collect information of each vulnerability from sources like CVE[20], CVEDetails[1], NVD[22], Openwall[23], SecurityFocus[24], Linux kernel Mailing List[25] and Git Linux kernel source tree[26]. Collected information included a brief description of the vulnerability, severity scores (usually CVSS[21] based), vulnerability types, vulnerable software and versions and references to advisories, confirmations, tracks, solutions, exploits and patches.

During the reproduction step, a subset of the vulnerabilities was selected to try or create a proof-of-concept (PoC) or exploit to reproduce them. Only those vulnerabilities with a low access complexity (according to CVSS score) or with a usable existing PoC or exploit were considered. A virtual machine with Debian 7.1.0 and a 3.12.0 vanilla kernel were used as a test system. Additionally, kernel dumps were configured and inspected to complement the verification step.

As an example, we show part of the analysis of the vulnerability CVE-2014-2523.



(a) CVEDetails (b) NVD

Fig. 3. CVE-2014-2523 information

Figure 3 shows the information recollected from CVEDetails and NVD websites. Both sites present the official description from the CVE database. NVD presents raw CVSS information, while CVEDetails present the same information but in a more descriptive way. Fortunately, the description is good enough to start the code analysis. Now we search any of the three mentioned functions (*dccp_new*, *dccp_packet* and *dccp_error*) and check for any use of a DCCP header pointer.

```
static bool dccp_new(struct nf_conn *ct, const struct sk_buff *skb,
                     unsigned int dataoff, unsigned int *timeouts)
{
    struct net *net = nf_ct_net(ct);
    struct dccp_net *dn;
    struct dccp_hdr _dh, *dh;
    const char *msg;
    u_int8_t state;

    dh = skb_header_pointer(skb, dataoff, sizeof(_dh), &dh);
    BUG_ON(dh == NULL);
```

Fig. 4. dccp_new vulnerable function

Table 2. CVE-2014-2523 Finding Summary

Property	Value
ID	CVE-2014-2523, 66279
Localization	net/netfilter/nf_conntrack_proto_dccp.c dccp_new (line 431), dccp_packet (line 489), dccp_error (line 580)
Description	Incorrect use of a DCCP header pointer, which allows remote attackers to cause a denial of service or possibly execute arbitrary code via a DCCP packet that triggers a call to the dccp_new, dccp_packet, or dccp_error function.
Cause	Buffer overflow trying to copy dccp header in a pointer address.
Prerequisites	Netfilter enabled in kernel configuration (disabled by default). conntrack and dccp_conntrack modules installed and running.
Remediation	Change &dh for &_dh in dccp_new, dccp_packet and dccp_error functions. (commit b22f5126a24b3b2f15448c3f2a254fc10cbc2b92)
Exploit	Non-existent. A PoC was created.
CVSS Severity	High (10.0) (AV:N/AC:L/Au:N/C:C/I:C/A:C)
DREAD Risk	High Priority (8.4) (DP:10/R:9/E:9/AU:6/D:8)

As we can see in Figure 4, *dh is the pointer we are looking for. This pointer is used to store the returned value of the function *skb_header_pointer*. As parameters, we send the buffer, a data offset, the size of the dccp header and the reference of the header; however, we are sending the reference of the pointer instead. If we take a look at the code of the *skb_header_pointer* function, it calls another function that copies the content to the referenced buffer. However, the reference received as a parameter is a pointer not the buffer. That means that if we receive a malformed DCCP packet, we could end overwriting memory on the stack causing a complete denial of service. Worth noticing that if we receive a special crafted packet and have no stack protection, we could end up executing some malicious code.

Finally we document the results of our analysis. As mentioned before we create a *finding summary*. Table 2 presents the finding summary for the vulnerability CVE-2014-2523.

4.2 Taxonomy

With our first analysis, we identified 7 facets:

– Cause of the vulnerability. Reflects the kind of error in the source code that caused the vulnerability.

- Location of the vulnerability. Refers to the location, in the source code, where the vulnerability is present and defined as subsystems of the kernel.
- Availability, integrity and confidentiality impact. Refer to the availability / integrity / confidentiality impact of a successfully exploited vulnerability.
- Access vector. Reflects how the vulnerability is exploited.
- Exploitability. This taxonomy classifies vulnerabilities according to the current state of exploit techniques and code availability.

The cause facet includes the next 5 categories:

- Validation Error: These vulnerabilities arise due to failures in responding to unexpected data or conditions.
- Synchronization and Timing Error: These are caused by the improper serialization of the sequences of processes or an error during a timing window between two operations.
- Resource Management Error: Vulnerabilities in this category are related to improper management of system resources.
- Exception Handling Error: These are caused by the improper response to the occurrence of exceptions.
- General Logic Error: These vulnerabilities are caused by a bad logic in the implementation. Errors like using the wrong operator, uninitialized variable, missing parameter, assigning instead of comparing and wrong operand order belong to this class.

The location facet has the next categories:

- System Call Interface (SCI): Layer that provides the means to perform function calls from user space into the kernel. Founded in /linux/kernel an /linux/arch.
- Process Management (PM). Focused on the execution of processes. Founded in /linux/kernel and /linux/arch.
- Virtual File System (VFS). Provides a common interface abstraction for file systems. Sources are founded in /linux/fs.
- Memory Management (MM). Sources founded in /linux/mm.
- Network Stack (NS). Sources founded in /linux/net.
- Arch-dependent code (Arch). Sources founded in /linux/arch.
- Device Drivers (DD). Sources founded in /linux/drivers. Part of the kernel that makes a particular hardware device usable.

The availability facet includes the next categories:

- Partial: There is reduced performance or interruptions in resources availability.
- Complete: There is a total shutdown of the kernel.

Confidentiality facet includes the following classes:

- None: There is no impact to the confidentiality of the system.

– Partial: There is considerable informational disclosure. Access to some system files is possible, but the attacker does not have control over what is obtained, or the scope of the loss is constrained.
– Complete: There is total information disclosure, resulting in all system files being revealed. The attacker is able to read all of the system's data (memory, files, etc.).

Integrity facet includes the following classes:

– None: There is no impact to the integrity of the system.
– Partial: Modification of some system files or information is possible, but the attacker does not have control over what can be modified, or the scope of what the attacker can affect is limited.
– Complete: There is a total compromise of system integrity. There is a complete loss of system protection, resulting in the entire system being compromised. The attacker is able to modify any files on the target system.

The access vector facet includes the next 3 categories:

– Local. Refers to vulnerabilities exploitable only having an interactive, local (shell) account that interfaces directly with the underlying OS.
– Local Network: These vulnerabilities require having access to either the broad-cast or collision domain of the vulnerable software.
– Remote. Refers to those vulnerabilities remotely exploitable.

And the exploitability facet includes the next four categories:

– Unproven: No exploit is publicly available or an exploit is entirely theoretical.
– Proof-of-Concept: Proof-of-concept exploit code or an attack demonstration that is not practical for most systems is available.
– Functional: Functional exploit code is available. The code works in most situations where the vulnerability exists.
– High: Either the vulnerability is exploitable by functional mobile autonomous code, or no exploit is required (manual trigger) and details are widely available.

4.3 Evaluation and Testing

For this phase, we only performed the evaluation and the test with our set of analyzed vulnerabilities. We classify our set of vulnerabilities on each of the described dimensions of our taxonomy. Tables 3 to 9 show the results of the classification.

During the classification, we identify the following problems on our taxonomy:

1. The facets' hierarchies need to be expanded. Subclasses are needed for a better description.
2. The location and access vector dimensions are ambiguous.

Because of that, we decided that at least another iteration of our process is necessary. The location dimension could be divided according to the source files and directories structure, rather than by subsystem. With the cause and access vector dimensions, hierarchical classes could be defined to avoid ambiguity.

Table 3. Vulnerabilities based in their cause

Type	Number	%
Validation Error	22	61.11
Synchronization and Timing Error	7	19.44
Resource Management Error	3	8.33
Exception Handling Error	2	5.56
General Logic Error	2	5.56

Table 4. Vulnerabilities based in their location

Location	Number	%
System Call Interface	3	8.33
Process Management	1	2.78
Virtual File System	3	8.33
Memory Management	3	8.33
Network Stack	14	38.89
Arch-dependent code	6	16.67
Device Driver	6	16.67

Table 5. Vulnerabilities by availability impact

Availability Impact	Number	%
Partial	1	2.78
Complete	35	97.22

Table 6. Vulnerabilities by confidentiality impact

Confidentiality Impact	Number	%
None	27	75
Partial	2	5.56
Complete	7	19.44

Table 7. Vulnerabilities by integrity impact

Integrity Impact	Number	%
None	27	75
Partial	2	5.56
Complete	7	19.44

Table 8. Vulnerabilities by existence of exploits

Existence of exploits	Number	%
Unproven	23	63.89
Proof-of-Concept	11	30.56
Functional	1	2.78
Complete	1	2.78

Table 9. Vulnerabilities by access vector

Access Vector	Number	%
Local	22	61.11
Local Network	6	16.67
Remote	8	22.22

5 Conclusion

The results of an analysis of DoS vulnerabilities on Linux, a new DoS vulnerabilities taxonomy and the classification of the analyzed vulnerabilities with the new taxonomy were presented. The analysis results will enable Linux kernel maintainers, developers and researchers to better understand denial-of-service vulnerabilities and prioritize their efforts according to the results presented. That way, better tools, techniques and defenses will be created. Differently from previous work, the scope of our taxonomy is DoS vulnerabilities in order to highlight the characteristics and provide more useful information with the classification.

References

1. CVE Details. Linux kernel, http://www.cvedetails.com/product/47/Linux-Linux-Kernel.html?vendor_id=33
2. Krsul, I.V.: Software vulnerability analysis. pp. 171. Purdue University (1998)
3. Lough, D.L.: A taxonomy of computer attacks with applications to wireless networks. pp. 348. Virginia Polytechnic Institute and State University (2001)
4. Aslam, T.: A Taxonomy of Security Faults in the UNIX Operating System. (1995)
5. Bazaz, A., Arthur, J.D.: Towards a Taxonomy of Vulnerabilities. ;In HICSS (2007)
6. Gegick, M., Williams, L.: Matching attack patterns to security vulnerabilities in software-intensive system designs. ;ACM SIGSOFT Software Engineering Notes(2005)1-7
7. Howard, J.D., Longstaff, T.A.: A common language for computer security incidents (1998)
8. Tsipenyuk, K., Chess, B., McGraw, G.: Seven Pernicious Kingdoms: A Taxonomy of Software Security Errors. ;IEEE Security & Privacy(2005)81-84
9. Hansman, S., Hunt, R.: A taxonomy of network and computer attacks. ;Computers & Security(2005)31-43
10. Killourhy, K.S., Maxion, R.A., Tan, K.M.C.: A Defense-Centric Taxonomy Based on Attack Manifestations. ;In DSN(2004)102-102
11. Landwehr, C.E.: Formal Models for Computer Security. ;ACM Comput. Surv.(1981)247-278
12. Kratkiewicz, K.: Evaluating Static Analysis Tools for Detecting Buffer Overflows in C Code. (2005)
13. Ahmad, N., Aljunid, S., Ab Manan, J.-l.: Taxonomy of C Overflow Vulnerabilities Attack. In: Zain, J., Wan Mohd, W., El-Qawasmeh, E. (eds.) Software Engineering and Computer Systems, vol. 180, pp. 376-390. Springer Berlin Heidelberg (2011)
14. Kratkiewicz, K., Lippmann, R.: A taxonomy of buffer overflows for evaluating static and dynamic software testing tools. Proceedings of Workshop on Software Security Assurance Tools, Techniques, and Metrics 500, 44 (2006)
15. Igure, V., Williams, R.: Taxonomies of attacks and vulnerabilities in computer systems. Communications Surveys & Tutorials, IEEE 10, 6-19 (2008)
16. Mirkovic, J., Reiher, P.: A taxonomy of DDoS attack and DDoS defense mechanisms. SIGCOMM Comput. Commun. Rev. 34, 39-53 (2004)
17. Ahmad, N.H., Aljunid, S.A., lail Ab Manan, J.: Understanding vulnerabilities by refining taxonomy. IAS, pp. 25-29. IEEE (2011)
18. Nickerson, R. C., Muntermann, J., Varshney, U., Isaac, H.: Taxonomy Development in information systems: Developing a taxonomy of mobile applications (2009)
19. Dowd, M., McDonald, J., Schuh, J.: The Art of Software Security Assessment: Identifying and Preventing Software Vulnerabilities. Addison-Wesley Professional (2006)
20. Common Vulnerabilities and Exposures, https://www.cve.mitre.org
21. Common Vulnerability Scoring System v2.0, http://www.first.org/cvss/cvss-guide
22. National Vulnerability Database, http://nvd.nist.gov
23. Openwall, http://www.openwall.com/
24. SecurityFocus, http://www.securityfocus.com
25. Linux kernel mailing list, https://lkml.org/
26. Linux kernel source tree, https://github.com/torvalds/linux

A Simple Web Interface for Inspecting, Navigating, and Invoking Methods on Java and C# Objects

Carlos R. Jaimez-González

Departamento de Tecnologías de la Información,
Universidad Autónoma Metropolitana, Unidad Cuajimalpa,
Av. Vasco de Quiroga 4871, Col. Santa Fe Cuajimalpa, C.P. 05300, México, D.F.
cjaimez@correo.cua.uam.mx

Abstract. In recent years there has been an increase in the development of different applications that rely on web browsers as their main interface to users and developers. This paper introduces a simple interface that allows inspecting and navigating Java and C# objects through any XML-aware web browser. The web browser interface also allows displaying and executing all the methods that belong to a specific Java or C# object that has been persisted on the system. The web browser interface is part of a framework for distributed object programming in Java and C#, called Web Objects in XML (WOX).

Keywords. Web Browser Interface, Java Objects, C# Objects, Object Navigation, Method Invocation, Web Objects in XML.

1 Introduction

Recently, there has been an increase in the development of different applications that rely on web browsers as their main interface to users and developers. This paper introduces a simple interface that allows inspecting and navigating Java and C# objects through any XML-aware web browser. The web browser interface also allows displaying and executing all the methods that belong to a specific Java or C# object that has been persisted on the system. The web browser interface is simple to use.

The web browser interface presented in this paper is part of a framework created for distributed object programming, called Web Objects in XML (WOX) [1], which allows building distributed systems; it uses XML as the representation for the objects and messages interchanged [2]; and provides both synchronous and asynchronous communication between clients and servers [3]. The WOX framework has special features; some of them taken from two paradigms used to construct distributed systems: the object-based paradigm and the web-based paradigm, where Java and C# objects can be persisted by distributed or local applications.

The rest of the paper is organized as follows. Related work is presented in section 2. An overview of the WOX framework is given in section 3. Section 4 describes how to inspect objects through the web browser interface. Section 5 provides an explanation on how to execute methods on specific objects that have been persisted on the

©L. Berthet, J. C. Chimal, E. A. Santos, E. Castillo
Advances in Computer Science
Research in computing Science 81, 2014 pp. 133-143

system. Section 6 discusses how to navigate objects through the web browser interface. Finally, conclusions and future work are provided in section 7.

2 Related Work

This section describes some related work used to browse objects. Although there are existing tools to browse objects, none of them is incorporated to a framework for distributed object programming, such as the one presented in this paper.

The Portable Explorer of Structured Objects (PESTO) [4], is an integrated user interface that supports browsing and querying of object databases. It allows users to navigate in a hypertext-like fashion, following the relationships that exist among objects. In addition, PESTO allows users to formulate object queries through an integrated query paradigm that presents querying as a natural extension of browsing, call it query-in-place. This interface can be configured to different object databases.

In [5] the authors propose an interactional operator dedicated to navigation through time, which would allow visualizing a snapshot of a collection of objects at a given instant, or detecting and examining changes within object states. They also study how this operator can be integrated with the two main types of interactions given in visual object database browsers: navigation within a collection of objects, and navigation between objects via their relationships, where users can explore and navigate the states of a set of related objects.

The authors of [6] worked more on the definition of a language rather than the development of a tool for navigation. They created XQBE (XQuery By Example), a visual query language for expressing a large subset of XQuery in a visual form, inspired by QBE, a relational language initially proposed as an alternative to SQL, which is supported by Microsoft Access. According to the hierarchical nature of XML, XQBE's main graphical elements are trees. One or more trees denote the documents assumed as query input, and one tree denotes the document produced by the query. Similar to QBE, trees are annotated so as to express selection predicates, joins, and the passing of information from the input trees to the output tree.

SOPView+ [7] is an object browser that supports navigation of a large database by changing the base object. The base object is an object which is a basis for navigation; forward navigation is provided for the reference paths ahead of the base object and backward navigation for the ones behind it. SOPView+ allows users to change the base object along the reference hierarchy among a number of database objects; this makes it possible for them to explore a large database until they find objects of their interest on the limited screen space, solving the screen real estate problem.

The CORBA Object Browser [8] is a web browser that can be used to directly invoke methods on CORBA Objects using a specifically designed URI scheme. The URI for an object not only identifies the object but may also optionally include the name of the method to be invoked on the object and the parameters required. It has been implemented by extending the HotJava browser. The browser has a mechanism for supporting access to CORBA Objects that run on a secure ORB through the

CORBA Object Browser. Accessing secure objects through the browser requires the authentication with the remote ORB and may also need secure communication.

3 Web Objects in XML

This section provides a brief overview of the WOX framework, which combines features of distributed object-based systems and distributed web-based systems. Some of the features of this framework are presented in the following paragraphs.

WOX uses URLs to identify uniquely remote objects, following the principles of the Representational State Transfer [9]. This is an important feature because all objects are uniquely identified by their URL and can be accessed anywhere on the Web, either through a Web browser or programmatically.

This framework uses an efficient and easy-to-use serializer, called the WOX serializer [2], which is the base of the framework to serialize objects, requests, and responses exchanged between clients and servers. This serializer is a stand-alone library based on XML, which is able to serialize Java and C# objects to XML and back again. One of its main features is the generation of standard XML for objects, which is language independent and allows reaching interoperability between different object-oriented programming languages. At the moment, applications written in the Java and C# programming languages can interoperate.

WOX has a set of standard and special operations supported on remote and local objects. These operations include the request for remote references, static method invocations (web service calls), instance method invocations, destruction of objects, request for copies, duplication of objects, update of objects, upload of objects, and asynchronous method invocations. Some of the operations are described in [1]. The mechanism used by WOX in a method invocation is shown in Figure 1.

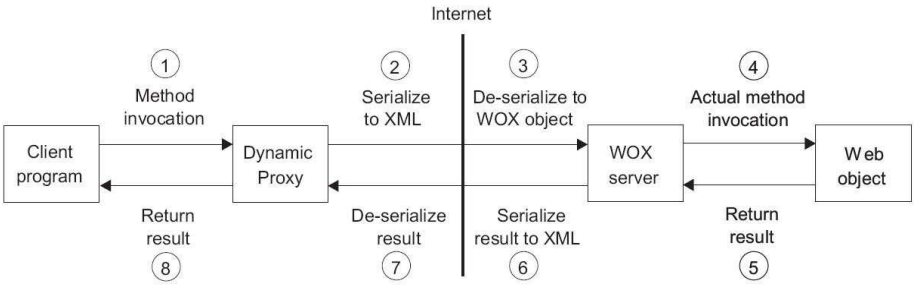


Fig. 1. A method invocation in WOX.

The detailed steps carried out in a method invocation in WOX are the following: 1) The WOX client program invokes a method on a remote reference (the way in which the client invokes a method on a remote reference is exactly the same as that on a local object, as far as the client program is concerned); 2) The WOX dynamic proxy takes the request, serializes it to XML, and sends it through the network to the WOX server; 3) The WOX server takes the request and de-serializes it to a WOX object; 4)

The WOX server loads the object and executes the method on it; 5) The result of the method invocation is returned to the WOX server; 6) The WOX server serializes the result to XML and either the real result or a reference to it is sent back to the client. The result is saved in the server in case a reference is sent back; 7) The WOX dynamic proxy receives the result and de-serializes it to the appropriate object (real object or remote reference); 8) The WOX dynamic proxy returns the result to the WOX client program.

From the WOX client program's point of view it just makes the method invocation and gets the result back in a transparent way. The WOX client libraries carry out the process of serializing the request and sending it to the WOX server; as well as receiving the result of the method invocation and de-serializing it.

The following sections introduce the web browser interface, which allows the inspection of objects through an XML-aware web browser, the execution of methods on objects, and the visualization of them with three different modes of operation: xml, html, and image. It can also navigate remote objects, which allows a client program retrieving child objects of a given root object, through their XML representation.

4 Inspecting Objects

Every Java or C# object that is persisted in the system, it is represented in XML, and can be inspected through a web browser by typing the URL that is assigned to it automatically by the system. Figure 2 shows a web browser with an example of the XML representation for an object of the *Lecturer* class. In order to inspect an object, the user types the object URL in the address bar of a web browser; for example: *http://carlosj:8080/WOXServer/WOXObject.jsp?id=622058786,* which is a URL representing an object with *id=622058786* on the system, which corresponds to an instance of the *Lecturer* class.

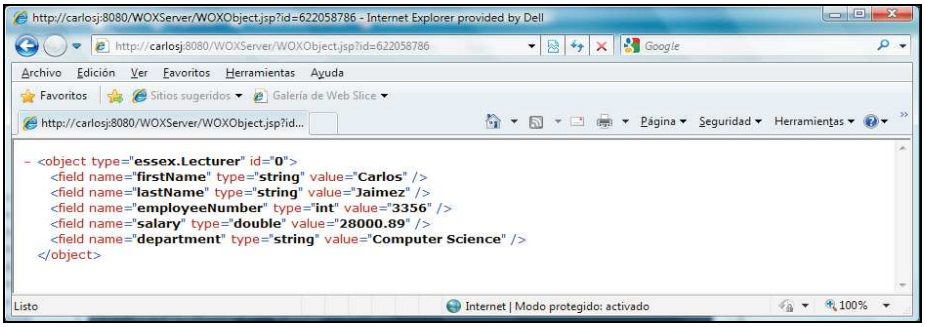


Fig. 2. Inspecting a Lecturer object through a web browser.

The XML representation of the *Lecturer* object is easy to understand, clean and compact. Every object is represented by an *object* XML element; and every field of an object is represented by a *field* XML element, with *name*, *type* and *value* attributes. The *Lecturer* object in Figure 1 has five attributes: *firstName* with *string* type;

lastName with *string* type; *employeeNumber* with *int* type; *salary* with *double* type; and *department* with *string* type. Figure 3 illustrates a diagram with the *Lecturer* class and an instance of it, which is represented in XML previously.

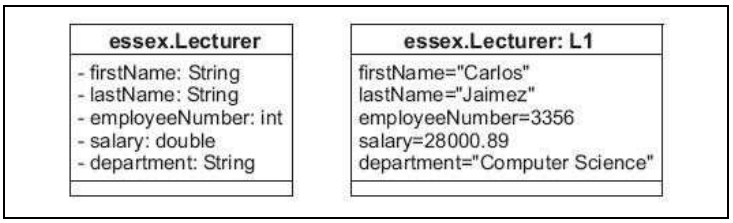


Fig. 3. The Lecturer class and an instance of it.

Alternatively, objects can be inspected through the server database, which holds all the objects that have been persisted on the system. Figure 4 shows an example of a server database, which can be accessed through the following URL: *http://carlosj:8080/WOXServer/WOXObjects.jsp*.

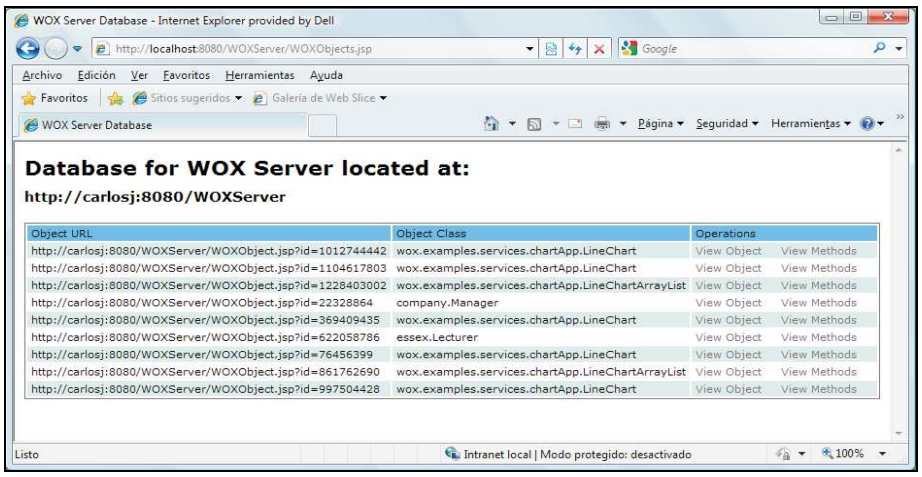


Fig. 4. A server database showing the objects persisted on the system.

The server database provides a table with three columns: the first column shows the URL of the object; the second column provides its implementation class; and the third column shows two hyperlinks. The *View Object* hyperlink can be used to request the specific object through its URL, and shows its XML representation on the web browser, as it was illustrated in Figure 1. The *View Methods* hyperlink is used to show the web browser interface with all the possible methods that can be invoked on the specific object. The invocation of methods is discussed in the following section.

The server database is an interface provided by the system, which allows access to persisted objects. This database can be seen as a directory for all the persisted objects,

where the user can request the object to inspect its XML representation or execute any of its methods available through the web browser interface.

5 Invoking Methods on Objects

Since every object can be identified uniquely through its URL, it is also possible to invoke methods on persistent objects through a web browser. As an example, the URL below shows an invocation of the *getDepartment* method, on an object of the *Lecturer* class, identified by the *622058786* object id: *http://carlosj:8080/WOXServer/WOXInvoke.jsp?objectId=622058786&method=getDepartment*.

The result of the method invocation is returned to the web browser as XML, which is the default mode of operation for the system, but it can also be returned as HTML (by specifying *mode=html* in the query string), in which case only the *string* would be returned. There is a special case in which the system can also return an image (by specifying *mode=image* in the query string), when the return type of a method is an array of bytes that represents an image.

Using a web browser, the system is also capable of invoking methods with parameters of primitive data types, but not with parameters of other data types, like user-defined classes. This way of operation through the web browser is similar to the way in which Apache Axis [10] allows invoking methods of classes. The main differences are that Axis, which is SOAP based [11], does not have the concept of an object, thus the methods are invoked as if they were static methods. Another important difference in this mode of operation is that Axis does not support package-qualified classes. In this respect, the system enables the invocation of methods on any web object (instance methods), and methods of any package-qualified class (static methods).

Alternatively, the system provides the browser interface with all the possible methods to invoke on a specific web object. This is similar to the browser front end for CORBA objects described in [8]. The web browser interface of the system can be accessed by typing the URL presented below, where it will present all the methods available for invocation on that object (in our example, it is a *Lecturer* object): *http://carlosj:8080/WOXServer/WOXInvoke.jsp?objectId=622058786*. Figure 5 shows this browser interface with three of the methods available for a *Lecturer* object. The web browser interface can also be accessed from the server database shown in Figure 4, by clicking the *View Methods* hyperlink of the desired object.

When clicking the *Invoke Method* button of the desired method, a query string is built internally with all the information needed for the method invocation. The request is sent to the system; the system will execute the method requested and generate an answer via an XML message with the result of the method invocation chosen; and finally this answer will be returned to the web browser (the return type of the method invocation is also specified in the web browser interface). The process of executing the method on the desired object is transparent to the user, since it is only needed to click the *Invoke Method* button to request the execution of the specific method, and wait for the answer.

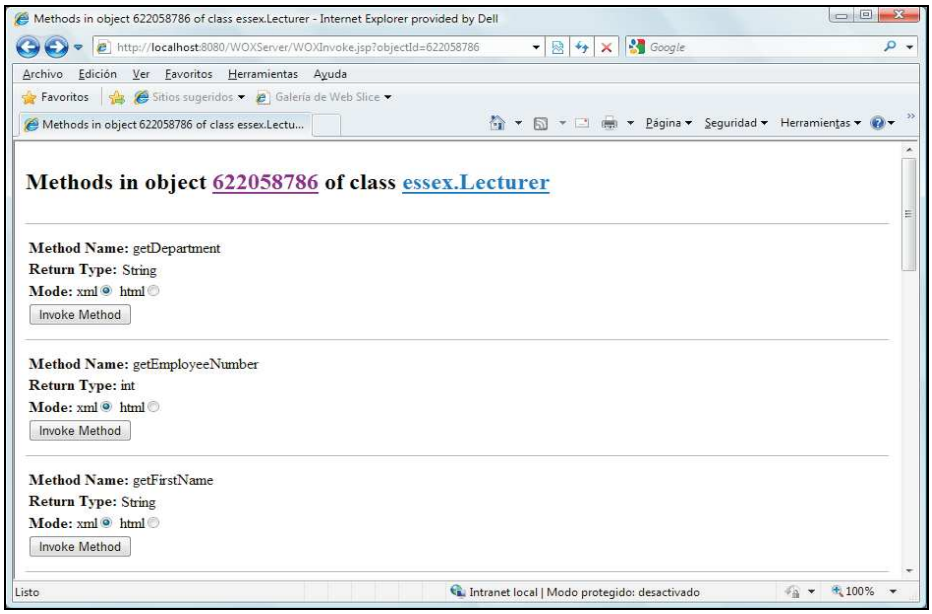


Fig. 5. Some methods available for a Lecturer object.

When the method to be invoked takes some parameters, the web browser interface provides the appropriate input boxes to enter the values to be sent with the method invocation. Figure 6 shows some other methods of a *Lecturer* object, which takes some parameters. It should be noticed that the input boxes shown in the web browser interface are generated automatically by the system, according to the data types of the parameters. Also notice that every parameter is named according to the order in which they appear in the method signature: *param1*, *param2*, etc.

In order to invoke the *setDepartment* method shown in Figure 6, the user must provide the *param1* parameter in the input box shown, which is of type *string*. When clicking the *Invoke Method* button of the *setDepartment* method, a query string is built with all the information needed for the method invocation on the specific object, which is the following: *http://carlosj:8080/WOXServer/WOXInvoke.jsp?objectId= 622058786&method=getDepartment&param1=Finance*.

The web browser interface provided by the system is a very convenient way of discovering and invoking methods on specific objects without needing a Java or C# client program.

6 Navigating Objects

The navigation of objects in the system allows a client program retrieving child objects of a persisted object. This is an important feature of this framework, which is very useful to retrieve only the child objects needed, instead of the entire object. This can also be carried out through a web browser. In this section we will use a different

object, since we want to show how to retrieve and navigate through an object which has child objects.

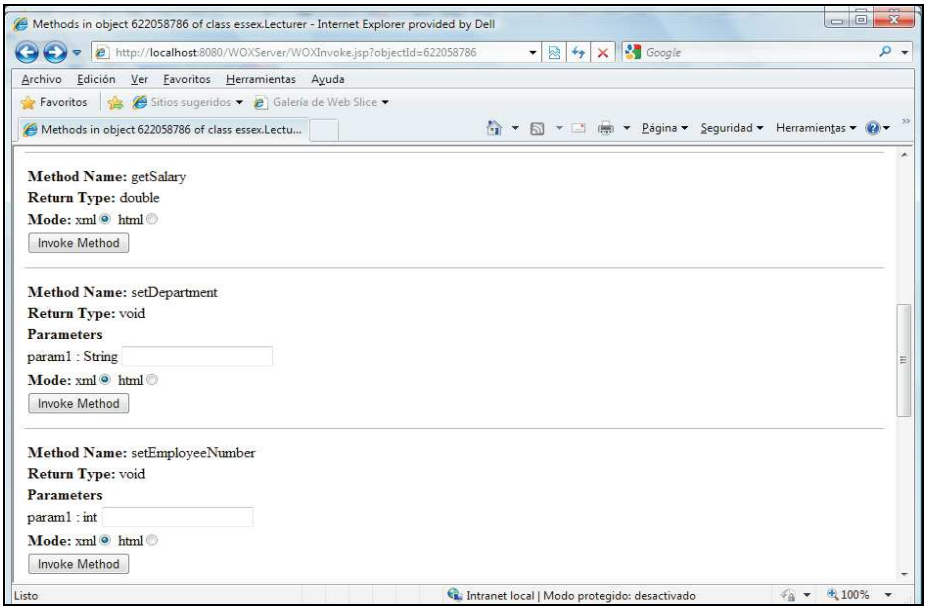


Fig. 6. Some methods with parameters available for a Lecturer object.

6.1 Retrieving an Object

This subsection shows how to retrieve an object. Assuming that the object to be retrieved can be accessed through the following URL: *http://carlosj:8080/WOXServer/WOXObject.jsp?objectId=622058786*. The XML representation of the object retrieved, which is a list of *Course* objects, is shown in Figure 7. It can be observed that the root object has three child objects identified by their *id* attributes with values 1, 2, and 3. The root object is identified by the *id* attribute with the value of 0.

6.2 Navigating an Object

The inner or child objects of a root object can be retrieved independently by specifying its *id* attribute in the URL of the specific object. For example, the following URL retrieves the *Course* object identified by *id="1"*: *http://carlosj:8080/WOXServer/WOXObject.jsp?objectId=622058786&id=1*.

The XML representation of the *Course* object after it has been retrieved is shown in Figure 8. It can be observed that only the child object specified has been retrieved. This is accomplished partly because of the XML representation of the objects, by using identifiers for every child object.

```
<object type="list" elementType="Object" length="3" id="0">
   <object type="Course" id="1">
      <field name="code" type="int" value="6756" />
      <field name="name" type="string" value="XML Technologies" />
      <field name="term" type="int" value="3" />
   </object>
   <object type="Course" id="2">
      <field name="code" type="int" value="9865" />
      <field name="name" type="string" value="O-O Programming" />
      <field name="term" type="int" value="2" />
   </object>
   <object type="Course" id="3">
      <field name="code" type="int" value="1134" />
      <field name="name" type="string" value="Web Programming" />
      <field name="term" type="int" value="2" />
   </object>
</object>
```

Fig. 7. XML representation of a list of Course objects.

Using this important feature, it is possible to retrieve only parts of a persisted object. This feature can be used through an XML-aware web browser where the object can be seen, but it can also be used programmatically from a Java or C# client application, using the framework for distributed object programming.

```
<object type="Course" id="1">
   <field name="code" type="int" value="6756" />
   <field name="name" type="string" value="XML Technologies" />
   <field name="term" type="int" value="3" />
</object>
```

Fig. 8. XML representation of a list of Course objects.

7 Conclusions and Future Work

This paper presented a simple web browser interface to inspect, navigate and invoke methods on Java and C# objects, which are persisted in the system. The web browser interface is part of Web Objects in XML (WOX), an interoperable framework for distributed object programming. It has been shown how objects can be inspected through a web browser; how to discover the methods to be invoked on them; and finally, how to navigate through objects in order to retrieve specific child objects. One of the key advantages of the web browser interface is that a user can inspect, navigate and execute methods on objects though any XML-aware web browser without needing any Java or C# program. The interface is intuitive and easy to use.

Further work is needed in order to allow the invocation of methods with parameters representing objects of user-defined classes, because at the moment parameters of methods can only be of primitive data types. Additional work is needed in this respect to provide an interface to upload this type of parameters. The fact of representing objects in XML is a great advantage, since they can be stored in simple XML files, which facilitates their management.

References

1. Jaimez-González, C., and Lucas, S., Implementing a State-based Application Using Web Objects in XML, In Proceedings of the 9th International Symposium on Distributed Objects, Middleware, and Applications (DOA 2007), Lecture Notes in Computer Science, Volume 4803/2007, pp. 577-594, Vilamoura, Algarve, Portugal, 25-30 November 2007.
2. Jaimez-González, C., Lucas, S., and López-Ornelas, E., Easy XML Serialization of C# and Java Objects. In Proceedings of the Balisage: The Markup Conference 2011, Balisage Series on Markup Technologies, Volume 7 (2011), doi:10.4242/BalisageVol7.Jaimez01, Montréal, Canada, 2-5 August 2011.
3. Jaimez-González, C., Lucas, S., and Luna-Ramírez, W., A Web Tool for Monitoring HTTP Asynchronous Method Invocations, In Proceedings of the 7th IEEE International Conference for Internet Technology and Secured Transactions (ICITST-2012), ISBN 978-1-908320-08-7, pp. 127-132, London, UK, 10-12 December 2012.
4. Carey, M., Haas, L., Maganty, V., Williams, J., PESTO: An Integrated Query/Browser for Object Databases, In Proceedings of the 22th International Conference on Very Large Data Bases, Mumbai, India, 3-6 September 1996.
5. Dumas, M., Daasi, C., Fauvet, M., Nigay, L., Pointwise Temporal Object Database Browsing, In Proceedings of the International Symposium on Objects and Databases, Sophia Antipolis, France, June 2000.
6. Braga, D., Capi, A., Ceri, S., XQBE (XQ uery B y E xample): A visual interface to the standard XML query language, ACM Transactions on Database Systems, 01/2005; 30(2):398-443, 2005.
7. Chang, S., Kim, H., SOPView+: an object browser which supports navigating database by changing base object, In Proceedings of the 21st International Conference on Computer Software and Applications Conference (COMPSAC 97), 1997.
8. Gupta D. Kumar, A. and P. Jalote. A browser front end for corba objects. In 10th International World Wide Web Conference, 2001.
9. Fielding, R., Architectural Styles and the Design of Network-based Software Architectures. Available at http://www.ics.uci.edu/ fielding/pubs/dissertation/top.htm, PhD thesis, USA, 2000. Last accessed in September 2014.
10. Apache Software Foundation. Web services - axis. Available at: http://ws.apache.org/axis/. Last access in September 2014.
11. World Wide Web Consortium. Latest soap version. Available at: http://www.w3.org/tr/soap/. Last access in September 2014.

Digital Signal Processing

Feature extraction module for word recognition on a mobile device based on discrete wavelet transform

Blanca E. Carvajal-Gámez[1], Erika Hernández Rubio[2], Francisco J. Hernández-Castañeda[2], Amilcar Meneses Viveros[3]

[1]Instituto Politécnico Nacional, Unidad Profesional Interdisciplinaria y Tecnología Avanzada, México D.F. México
[2]Instituto Politécnico Nacional, SEPI-ESCOM, México D.F. México
[3]CINVESTAV-IPN, Departamento de Computación, México D.F. México
[1,2]{ehernandezru, becarvajal}@ipn.mx, [2]hcasfj@gmail.com
[3]ameneses@cs.cinvestav.mx

Abstract. The voice is a natural means of interaction between users and mobile devices. Currently, most of the applications based on speech recognition for mobile devices performs searches on large databases through web services without strict voice recognition. These web services require too many resources of storage and processing, which creates a dependency between the mobile device and the communications network. In this paper a word recognition algorithm is presented to perform speech recognition. The processing algorithm is performed in the mobile device. The main contribution is the recognition of words, not just limited to syllables. Word recognition is performed by extracting characteristics through discrete wavelet transform- Haar, to calculate the estimate of the variation of the field of sampling and finally we applied fuzzy logic.

Keywords.- speech recognition, mobile devices, discrete wavelet transform, fuzzy logic.

1. Introduction

The ability to recognize human speech has been an area of interest mainly to the large variety of applications that can be performed. One of the priorities in the development of mobile device technology is to improve the quality and capacity for speech recognition. The speech recognition allows mobile devices adapt the voice information in a comprehensible way, which means the identification and understanding of the information received[1].

One objective of speech recognition systems is to facilitate communication between the user and the device, through the development of applications. There are applications for control systems, through spoken commands, utilizing a voice user interface (VUI), useful for controlling multiple devices, especially for people with certain physical disabilities or working with multimodal interfaces [2]. Other applications allow the user iteration with lexical recognition systems, predefined or online. In this paper, an adaptive VUI [3] mobile device is proposed, to identify words in a noisy environment.

©L. Berthet, J. C. Chimal, E. A. Santos, E. Castillo
Advances in Computer Science
Research in computing Science 81, 2014 pp. 145-155

Notwithstanding the recent development of methods for speech recognition based in a data base (Google speech recognition API) [4] and synthesis [4], implementations have been performed on personal computers and through remote servers, still have some unresolved needs of embedded systems for mobile devices, autonomous systems, offline systems, control systems, interactive voice, computer equipment for people with physical disabilities. The recognition of large vocabulary of the human speech requires complicated algorithms with big amounts of computations and large memory spaces [1-4]. This results in high energy consumption, greater recognition time and increased error rate. In this paper we use of the discrete wavelet transform (DWT), for extraction of the main characteristics of the voice signal, is proposed in addition to providing compression of the voice signal, resulting in reduced energy consumption, computational complexity and eliminates dependence on the network with an Internet connection. The proposed algorithm with the DWT, decomposes the signals in different frequency components, it possible to locate a specific vibration signal [5]. The DWT can be used for a wide variety of signal processing tasks, such as compression, noise elimination and enhancement of recorded speech [5]. DWT Haar is considered due to the facility in performing its calculations and characteristics relating to the preservation, compaction and redistribution of energy, all without altering the original signal [5]. After applying the compression and obtaining the spectral content of the audio signal with the DWT; continues for extracting characteristics, such as: the variance, standard deviation and the average, and the recognition result is obtained by means of a logic fuzzy that takes as input the result of the application of the DWT-Haar. A method of speech recognition for mobile devices offline for uncontrolled environments, as well as a system to reduce dependence on the internet is proposed. This system of speech recognition is installed on a smartphone with Android 4.0 operating system, with a RAM of 512 MB and a processor Qualcomm MSM7225A @ 800 MHz-1256 DMIPS, with which superior results were obtained at 70% for noisy environments these results are shown in the fourth section. This paper is organized as follows: the second section presents some recent works, the third section presents the architecture and the speech recognition system proposed, the fourth section presents the speech recognition tests performed, and an analysis of the results and finally, conclusions are presented.

2. Related work

According to [1], speech recognition is the ability of a machine or other device to identify words or phrases and convert these in action.

There are two types of speech recognition systems:

I. recognition systems based on speech recognition methods operate directly on a voice terminal.
II. recognition systems based on methods of speech recognition voice operated from a server.

Those that use a server to transmit the sound signal or voice characteristics to the server, which runs on a search engine and returns the speech recognition to the device, it runs so online. The speech recognition systems operate on a server has a

search engine of the word possible matches found, this type of recognition systems is restricted to small vocabularies of certain areas or geographic areas[1].

Placing a speech recognition system on a mobile device and is connected to the server to continually search for the word causes the energy consumption is higher. Currently, the way speech recognition for mobile devices is working through services on the network. The speech recognition is essentially an integration of automatic speech recognition (ASR) and search for the words that match the voice in a database located on a server [6-7]. To improve search precision, some voice search systems also include a module for natural language understanding (NLU) to analyze the output of the ASR in significant segments [6]. Services of speech recognition in network resource-intensive storage and processing[8-10]. An example of how these resources are used in [11] where some of the important features for the development of the Voice Search application are listed by Google as it is the power of the database and the schema for the generation of the structure recognition of patterns of language you lie to the specific case of the creation of the modeling language together its relevant database is shown in [12], this for the Mandarin-Chinese language by Google.

Finally in [13] a system for speech recognition to the Japanese language is done using the Google database. The use of these resources creates a dependency between the mobile device and the communication network as well as has some other problems such as vulnerability to noise, high computational complexity, high memory consumption, high energy consumption and unfriendly interfaces to user [8-10].

3. Proposed method

The most significant problems in speech recognition systems are related to the individuality of the human voice (such as age or gender to name a few), dialect, speaking rate, the context of phonics, noise background, the characteristics of voice acquisition device (microphone), and the directional characteristics of the source speech signal, among others [1-3]. All these problems are considered in the proposed method, because the objective of this research is that the speech recognition is inherent to the environment and the people who use the application. The speech recognition module, based on the DWT-Haar, comprises three main blocks: encoding and compression, feature extraction and recognition, as shown in Figure 1.

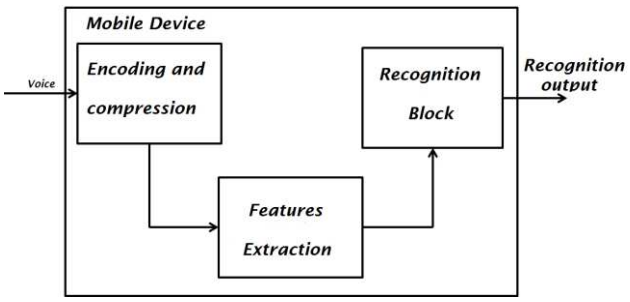


Fig. 1 – Block diagram of the proposed method.

The following subsections explain each of the blocks of the proposed module.

3.1 Block for enconding and compression

In this block, two vectors are obtained: The approximation vector (AV) and the fluctuations vector (FV).

These vectors are obtained when DWT-Haar is applied in the original speech signal. The size of the vectors AV and FV is half the size of the original speech vector. The AV vector contains the low frequency components of the voice signal. The vector FV contains high frequency components of the speech original signal. For this work the AV was chosen, because this vector has the largest amount of information about the speech original signal. To encode speech signal through native methods of recording, signal compression is obtained through the AV with WAV format. Compression has a rate of 22050 samples per second with 16 bits per sample without encoding pulse code modulation (PCM). In the Figure 2, the wavelet decomposition of the original signal is shown in the vector AV $y_{Lo\_D[w/2]}$, and the vector FV, $y_{Hi\_D[w/2]}$. Where $w/2$, is half of the speech signal original vector length.

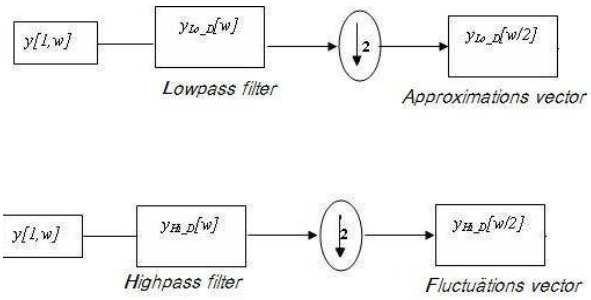


Fig. 2 – Block diagram of wavelet decomposition to the speech original signal.

With the wavelet decomposition of the speech signal and using the vectors AV-FV, we can see the spectral content of the signal, as shown in Figure 3.

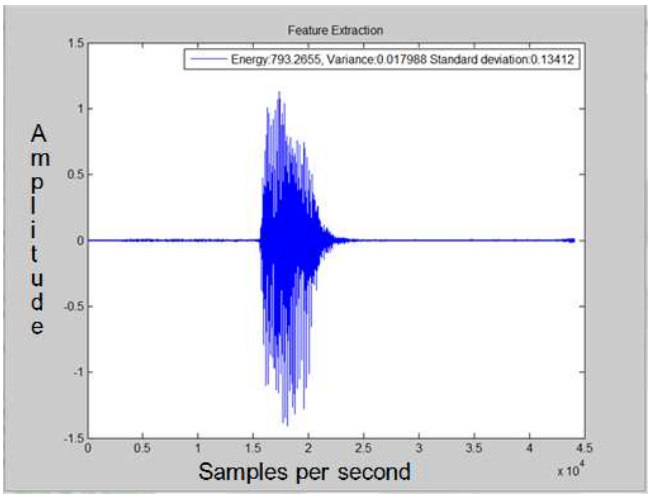


Fig. 3 – Approximations sub-vector obtained of the speech original signal.

In Figure 3, the plotting of the AV vector, obtained from DWT-Haar decomposition of the speech signal is displayed. This graph shows the reduction of 44,100 samples per second (standard WAV file) to 22,050 samples per second. In this vector, has the highest energy input speech signal therefore has the highest content of key features to extract the spectral content of the speech signal and is found in the word contained.

3.2 Block for extracting features

This block obtains the corresponding features of each input speech signal. This extraction is performed in the AV obtained in the previous block. Acquired characteristics are energy Eq. (1), the standard deviation Eq. (2), variance Eq. (3) and the center frequency of the speech signal in the AV.

$$E[y_{Lo\_D}[n]] = \sum_{m=1}^{n} \left| y_{Lo\_D}[m] \right|^2 \tag{1}$$

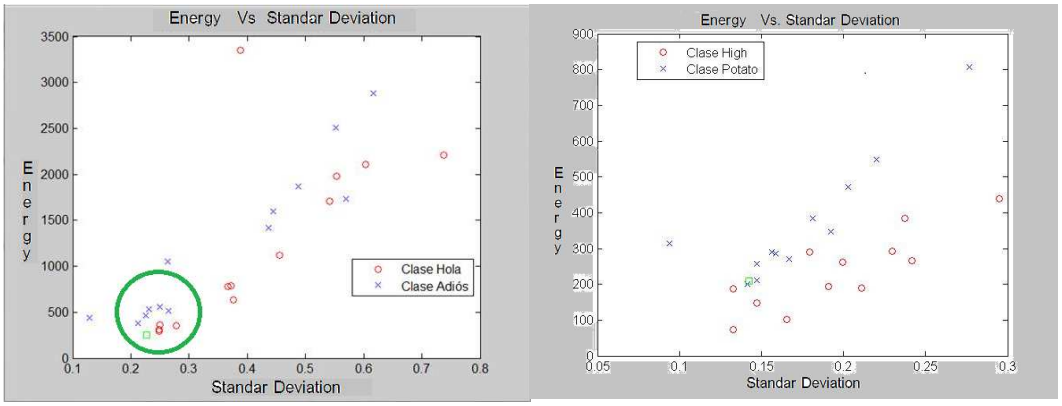
$$\sigma_{y_{Lo\_D}} = \sqrt{\sum_{m=1}^{n} \left(y_{(Lo\_D)_m} - y_{Lo\_D} \right)^2 / n} \tag{2}$$

$$\sigma^2_{y_{Lo\_D}} = \sum_{m=1}^{n} \left(y_{(Lo\_D)_m} - y_{Lo\_D} \right)^2 / n \tag{3}$$

Where $y_{Lo\_D} = \sum_{m=1}^{n} y_{Lo\_D} / n$ represent de mean value of AV, and $n = w/2$.

3.3 Block to recongnition

This block is determined by the spoken word. The words used are suggested: "Hola" and "Adios", and also "High" and "Potato". The words "Hola" and "Adios" in particular were chosen in order to show that the system works with high statistical dependence as shown in Figure 4 a) , contained in the green circle overlapping of these two words is observed.



a) b)

Fig. 4 – Graphic of energy vs. standard deviation, a) words “Hola” y “Adios”, b) "High" and "Potato”.

Unlike the words "High" and "Potato", which in Figure 4 b), the statistical independence is observed. This task is accomplished by using speech recognition. The entries in this block are the characteristics of the speech signal.

Fuzzy logic was used for word recognition in speech signal. This technique allows determining whether a spoken word in the set of speech signal features vector. Fuzzy logic is performed using the membership functions [14] for each of the features extracted from the equations (1), (2), (3) and (4). The membership function indicates the degree to which each element of a given universe belongs to a set. If the set is crisp, the membership function (characteristic function) take the values {0,1}, while if the set is blurred, it will take in the interval [0,1]. If the result of the membership function is equal to 0, then the element is not in the set. In contrast, if the result of the membership function is 1, then the element belongs to the set completely[14]. Gaussian membership, Eq. (4), is used for the purpose of this word recognizer in the speech signal by using fuzzy logic. The Gaussian membership is specified by two parameters $\{c,\sigma\}$ to determine the boundaries of the speech signal, and also to determine where the greatest amount of information is presented in the spectral content of the word to be identified.

$$Gaussian(x;c,\sigma) = e^{-\frac{1}{2}(\frac{x-c}{\sigma})^2} \quad (4)$$

The Gaussian function is determined by the values they take σ and *c*. Where*c* represents the center of the function and σ is the standard deviation (2). For this case, *c* is the mean value, and each value is the mean average of each standard sample, and σ is the standard deviation (2) of each test pattern.

4. Tests and results

The module for speech recognition embedded was implemented on a mobile device. This module identifies words: "Hola" and "Adiós", and also "High" and "Potato".
The speech recognition system was tested in a smartphone with Android 4.0 operating system, 512 MB RAM and Qualcomm MSM7225A @ 800 MHz-1256 DMIPS processor. A population with diverse gender and age were chosen for testing. Samples are shaped with a total population of 12 people, 6 men and 6 women. The ages are classified as follows. People over 50 years old: 3. People between 30 and 50 years old: 3. And people between 20 and 30 years old: 6.

To execute the embedded module, the audio is acquired through the microphone of the mobile device. The speech files that were used have different environmental conditions, because they want to test the robustness and accuracy of the identification of the word within the audio file. The module was also tested in different languages.

Tables 1 and 2 show the results obtained from the algorithm for the embedded module. In these tables you can see the quantitative results of this algorithm. Also a comparison between the signal acquired and processed audio signal is shown.

Table 1. Results obtained by the embedded module to identify "*Hola*" and "*Adiós*"

	Embedded module			
Age range	**Processing time (seg)**	**Sp(%)**	**Se(%)**	**ACC(%)**
age>50	0.109	69.35	93.33	71.08
age>=30&&age<=50	0.116	71.87	96.66	66.66
age>=20&&age<30	0.106	69.49	100	70.55
Average	0.110	70.23	96.66	69.43

Table 2. Results obtained by the embedded module to identify "*High*" and "Potato"

	Embedded module			
Age range	**Processing time (seg)**	**Sp(%)**	**Se(%)**	**ACC(%)**
age>50	0.107	68.25	98.90	73.34
age>=30&&age<=50	0.120	64.82	97.56	72.22
age>=20&&age<30	0.120	63.22	100	73.33
Average	0.115	65.43	98.82	72.96

Performance results are calculated from the speech signal processing through the mobile device. To test the performance of the proposed method, we consider four cases: two for correct classifications and two for misclassification. The classifications

are: true positive (TP), false positive (FP), false negative (FN) and true negative (TN). By using these different measures of performance metrics as the following relation is obtained [15]:

$$Specificity = TN/(TN + FP) \tag{5}$$

$$Sensitivity = TP/(TP + FN) \tag{6}$$

$$Precision = (TP + TN)/samplesofspeechsignal \tag{7}$$

Specificity (Sp) is the ability to detect samples that do not correspond to the audio signal. The sensitivity (Se) reflects the ability of an algorithm to detect the sample audio signal. Accuracy (ACC) measures the ratio between the total number of correctly classified samples (sum of true positives and true negatives) by the number of samples of audio signal[15]. The positive predictive value, and accuracy rate, gives the proportion of samples of the audio signal identified which are true [12,16]. That is, the probability that a sample of the audio signal is identified as true positive. From the results of tests concentrated in Tables 1 and 2, we note the following. The module embedded on the mobile device detects up to 70% of search word (Sp) and up to 96.66% for detecting those who have the search word in the audio file (Se) for the Spanish language. For English language, the embedded module detects up to 65.43% the search word (Sp) and up to 98.82% for detecting those who have the search word in the audio file (Se). In the case of the Spanish language, an accuracy greater than 69.43% is obtained at a time of 0.110 seconds. And for English, an accuracy of 72.96% is obtained in a time of 0.115 seconds. The range of ages, regardless of gender, who presented the best accuracy for word recognition within the mobile device is above 50 years old, for the Spanish language, Table 1. In the case of English language, the same behavior was presented. The greatest accuracy was obtained with the elderly 50 years old, Table 2.

In the case of the Spanish words that begin or end with the same phoneme, such as cases that were used in this research as "Hola" and "Adiós" as well as the noisy environment of the cell itself and the age range of the people, causing the rate of Sp, and Acc is, decrease its performance.

In [4], a module for recognizing isolated words, in real time, on a mobile device is presented. In this study conducted with a sample population of 10 people in total. This population consists of 5 men and 5 women, and the age range of the members of the population is not mentioned. Replays of the audio signals on the mobile performed 3 times to create an average error. The tests in this work were attacked with white noise with SNR, Eq. (8), between 15 and 30 db[4].

$$SNR = \frac{\bar{y}_{Lo\_D}}{y_{Lo\_D}} \tag{8}$$

The SNR has a uniform distribution, so that altering the signal audio not affected and a substantial modification of this signal is taken. This is different to present audio signal to ambient noise, because in this case the signal is affected by impulsive noise. And this kind of noise no predictable distribution, generating in certain sections of

audio are altered significantly. The results presented in [4], are returned in a time of 11.61 ms and with an average of 61.4% in the worst case, when the signal is changed to white noise. And an average of 90.2% is obtained when the audio file is in a controlled environment.

5. Conclusion

The proposal presented in this research for isolated word recognition in a mobile device in uncontrolled environments gives yields higher than 70% in less than the time 0.120 seconds. This embedded word recognizer module requires no prior training or generation of a dictionary as those currently commercially. Also the so-embedded module works offline, ie; not require a connection to the network in order to perform their job recognition. This streamlines its use and management, adding portability and the generation of an App to be used as a tool in voice commands or support systems in any treatment such as Luria tests. We note also that being a working embedded system offline use so the battery is not as affected in the performance of this. The word recognition system achieves work for any genre and any age group not presenting any difficulty, to be altered or amended by voice acuity or severity of the tone of speech signal for the gender of the user, as well as the possible echo the voice generated by age.

Concluding finally that although in some cases the performance is not expected, than the results shown in [4], embedded system mounted on an FPGA, where the processing is done faster and transparent manner.

Aknowledgment

The work is supported by Instituto Politecnico Nacional (IPN), Mexico.

References

1. Husnjak S.; Perakovic D.; Jovovic I.: Possibilities of using Speech Recognition Systems of Smart Terminal Devices in Traffic Environment; *Procedia Engineering*, vol. 69, p.p: 778 – 787, 2014.

2. Oviatt S., Julie A. Jacko, *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications*, CRC-Press, p.p: 286-304 , 2003.

3. Ons B.; GemmekeJort F.; Van Hamme H.: Fast vocabulary acquisition in an NMF-based self-learning vocal user interface, *Computer Speech and Language*, vol. 28, p.p: 997–1017, 2014.

4. Sledevic T.; Tamulevicius G.; Navakauskas D.: Upgrading FPGA Implementation of Isolated Word Recognition System for a Real-Time Operation, *ELEKTRONIKA IR ELEKTROTECHNIKA*, vol. 19, no.10, p.p: 1-6, 2013.

5. Walker,J. S., University of Wisconsin, *A Primer on WAVELETS and Their Scientific Applications*, Chapman and Hall/CRC, Second Edition. 2008.

6. Feng J.; Johnston M.; Bangalore, S., Speech and Multimodal Interaction in Mobile Search, *Signal Processing Magazine*, vol. 28, no. 4, p.p:40-49, 2011.

7. Jeong, H.D.J.; Sang-Kug Y.; Jiyoung L.; Ilsun Y.; Wooseok H.; Hee-Kyoung S., A Remote Computer Control System Using Speech Recognition Technologies of Mobile Devices, *Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS)*, p.p: 595-600, 2013.

8. Cohen, J.: Embedded speech recognition applications in mobile phones: Status, trends, and challenges, *Acoustics, Speech and Signal Processing*, IEEE International Conference, p.p: 5352-5355, 2008.

9. Marshall J.; Tennent P.: Mobile interaction does not exist, *Extended Abstracts on Human Factors in Computing Systems*, p.p: 2069-2078, 2013.

10. Flynn R.; Jones E.: Speech Enhancement for Distributed Speech Recognition in Mobile Devices, *Consumer Electronics*, International Conference, p.p: 1-2, 2008.

11. Schalkwyk J.; Beeferman D.; Beaufays F.; Byrne B.; Chelba C.; Cohen M.; Kamvar M.; Strope B.; Google Search by Voice: A case study, Weinstein, A. Visions of Speech: Exploring New Voice Apps in Mobile Environments, *Springer* 2010.

12. Shan J.; Wu G.; Hu Z.; Tang X.; Jansche M.; Moreno P.: Search by Voice in Mandarin Chinese, *Interspeech*, p.p: 354-357, 2010.

13. Shimada T.; Nisimura R.; Tanaka M.; Kawahara H.; Irino T.: Developing a method to build Japanese speech recognition system based on 3-gram language model expansion with Google database, *Conference Anthology IEEE* , p.p: 1-6, 2013.

14. Cavus N.: The evaluation of Learning Management Systems using an artificial intelligence fuzzy logic algorithm, *Advances in engineering software*, vol. 41, no. 2, p.p: 248-254, 2010.

15. Muñoz Pérez C., Cabrera Padilla D., Carvajal Gámez B.E., Gallegos Funes F.J., Gendron D., Segmentación automática en imágenes RGB aplicando la técnica Fuzzy C-means morfología matemática para la ayuda de la foto-identificación de cetáceos, *COMIA 2014*, in press

16. Kumar A.; Tewari A.; Horrigan S.; Kam M.; Metze F., Canny J.: Rethinking Speech Recognition on Mobile Devices, *ACM*, p.p: 1–6, 2011.

SECOND ORDER IDENTIFICATION FILTER

J. J. Medel J [1], M. T. Zagaceta A[2], K. A. Aguilar C[3]

[1, 3] Computing Research Center Av. 100 Metres, St. Venus S/N, Col. Nueva Industrial Vallejo, CP. 07738
[2] Mechanical and Electrical Engineering School Av. De las Granjas N. 682 Col. Santa Catarina CP 02250.

Abstract. The identification process shows the internal dynamic states based on a reference system commonly known as a black box scheme, built on the transition and gain functions, and an innovation process. Unfortunately, in this sense, the exponential transition function considers the unknown internal parameters. This means that the internal states description does not operate with these conditions considering the internal dynamic gains inaccessible. Against this difficult there is an approximation using the estimation technique. This paper presents estimation for a single input - single output (SISO) model with two delays known as a second order system. This is a special technique which describes the black box internal gains, allowing the transition function to have a sense of identification. The estimator built on the second probability moment form has an adaptive strategy allowing a sufficient convergence rate into identification results. The filter is integrated with two actions, estimation and internal states description, seeking a good convergence level considering the gradient description. Therefore, the theoretical result demonstrates the filter strategies.

Keywords: Digital Filter, Estimation, Functional Error, Identification, Stochastic Gradient, Reference Model.

1. Introduction

A physical system must be performed as a model mathematically validated with different processes. The difference between the real output system and its mathematical output representation tends to be a predefined bounded region. The model as a black box scheme establishes a relationship between the output and input signals system. In particular, the internal states are non-observables and are related mathematically between the input and output signals through the transfer function.

The system viewed as a black box permits the transfer function without access to their internal dynamics in a direct form. In this case the filter process is necessary describing the states indirectly; i.e., the identification is a technique describing the internal states [1-2]. This system commonly gives the output online without knowing what happens inside at a specific time. Therefore, the techniques were developed and applied into the identification filter structure [3]. The internal dynamics description

©L. Berthet, J. C. Chimal, E. A. Santos, E. Castillo
Advances in Computer Science
Research in computing Science 81, 2014 pp. 155-163

was based on the early work of identification depending on the transition function [4], which in its simplest form obtains the primitive equation including the internal gain [5]. The digital filter identification can be seen illustratively in Figure 1.

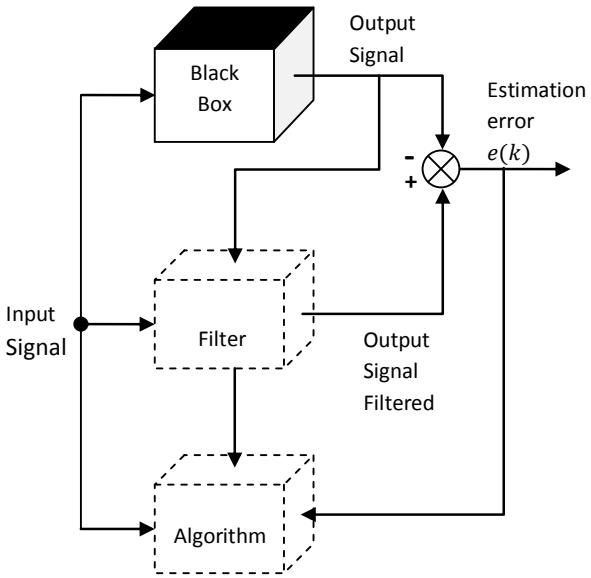


Figure. 1. Black Box and Digital Filter.

The black box system evolution is achieved through interpolation methods approximating the response to a real output system [6]. The parameter function plays an important role in the convergence rate considering it affects the model evolutions [7].

In control theory, equations are commonly used in the states space meaning the mathematical model depends on internal states and parameters [8]. In spite of this, internal unknown conditions are described using combined estimation and identification techniques [9]. The output velocity state changes with bounded movements, so that the model proposed corresponds to a second order with two delay states [10]. Therefore, the parameter estimation in state spaces in discrete form is applied into an identification model [10]. Once estimated and adapted to an identification technique, the evolution adjusts the estimation results converging the output identification filter to the desired output answer [11].

The estimation technique is classified into recursive and non-recursive. A Non-recursive technique corresponds to a final parameter vector. A recursive techniques process evaluates the evolution parameters through time [12]. Once parameters are obtained, the transition function is built [13].

The internal dynamics reference system in simple form considers the input and output bounded signals viewed as linear and stationary in the probability sense [14],

with unknown parameters [15, 16]. The traditional identification process is not feasible if the gain used for the transition function is inside the black box [17].

The transition function being applied in the adaptive process is based on: 1) the initial conditions of the transition function. 2) The second probability identification error moment form. 3) A criterion dynamically modifies the gain.

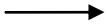
The identification error was obtained by the difference between the reference system response and the filter. To achieve the identification, it is necessary to know the transition role and parameters through the second probability moment. Once internal parameters are estimated, the transition function is built to be applied in identifying the internal states. The identifier signal and the functional error built on the basis of the second probability moment correct the estimation dynamically, making the interaction lead to the filter in the convergence region.

The estimation and identification are related by the functional error, obtaining a dynamic digital filter. According to the following order, the retrieved estimation is a gradient stochastic type accessing the system output. The identification is described by the recursive structure and also, the functional error. The estimation and identification findings are presented. In general, the importance of the transition in identification is solved using the estimation as an adaptive procedure.

2. Theoretical Analysis

The estimation describes the black box internal parameters system, but is it possible to estimate parameters that affect the identification filter? The answer is yes, because the transition function affects the convergence between the identification and reference signal. Therefore, the estimation would be part of the identification considering the transition function built with estimation parameters. The adaptive process is built considering that the estimation is based on functional identification error affecting the gains in direct form as shown in Figure 2 using the symbols of Table 1, described in finite differences.

Table 1. Black box symbols

	Input, From input to output, Output,
z^{-1}	Signal delay , $x_{k-1} := x_k z^{-1}$
	Algebraic adding signals
	Algebraic multiplicity signals

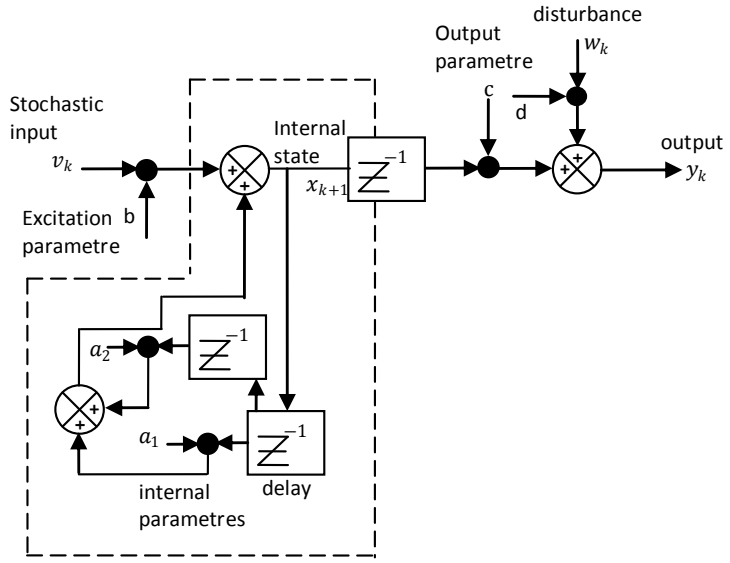


Figure. 2. Filter Estimation Scheme.

Figure 2 considers the output state space (v_k) with external perturbation (w_k). Inside the box with a dashed line is known as dynamic internal parameters described in state space recursive form.

The reference system according to [18] has a recursive form (1) in which the state space according to [19-21] is a function of $x_k = f(a_1, a_2, x_{k-1}, x_{k-2}, b, v_k)$; with output $y_k = f(c, x_k, d, w_k)$, as shown in Figure 2, where $\{w_k \coloneqq w(k), k \geq 0\}$ is a stochastic description bounded by a Gaussian process $N(\mu = k_{w_k}, \sigma^2_{w_k} < \infty)$. , and $\{v_k \coloneqq v(k), k \geq 0\}$ with $N(\mu = k_v, \sigma^2_w < \infty)$.

In Figure 2, the identification filter requires knowing the estimated internal parameters. Then the identification filter includes the estimation technique implemented as a system.

Theorem 1 (Adaptive Identification). Consider a black box system described by a second-order stochastic model expressed in finite differences (1)

$$y_k = c_k x_k + w_k \text{ with an internal state } x_k = a_1 x_{k-1} + a_2 x_{k-2} + b v_k . \quad (1)$$

Bounded in time evolution $\tau_k < \infty$ according to [22] and $f_{\max_k}$ is the frequency representing either bounded $f_{\max_k} < \infty$ [23] with respect to $\{w(k), k \geq 0\}$ and $\{v(k), k \geq 0\}$ such as discrete stochastic processes described by $N(\mu_w = k_w, \sigma_w^2 < \infty),\ k_w \in \Re_+$, $N(\mu_v = k_v, \sigma_w^2 < \infty),\ k_v \in \Re_+$, and satisfying $x_k, x_{k+1},\ y_k \in \Re_+^{[n\times 1],k}$; with parameters $a_1,\ a_2 \in \Re_+$ based on an error identification and its variance with respect to [20-23] and $\hat{y}_k$ described in (2).

$$\hat{y}_k = \hat{a}_1 y_{k-1} + \hat{a}_2 y_{k-2} + \hat{d}\hat{w}_k \tag{2}$$

with

$$\hat{a}_{1,k} = \frac{E\{\hat{y}_k y_{k-1}\})}{E\{y_{k-1}^2\}},\ \hat{a}_{2,k} = \frac{E\{\hat{y}_k y_{k-2}\})}{E\{y_{k-2}^2\}}$$

Proof (Theorem 1): Considering the recursive functional error in (3).

$$J_k = \tfrac{1}{k}(E\{e_k^2\} + (k-1)J_{k-1}). \tag{3}$$

The identification error described in (4):

$$e_k := y_k - \hat{y}_k \tag{4}$$

According to the system (1), the output signal is described in (5).

$$y_k = c(a_1 x_{k-1} + a_{2,} x_{k-2} + b v_{k-1}) + d w_k \tag{5}$$

With one and two delays, the output signal is (6):

$$y_{k-1} = c x_{k-1} + d w_{k-1}\ ,\ y_{k-2} = c x_{k-2} + d w_{k-2} \tag{6}$$

The unobservable states for y_{k-1} and y_{k-2} are solved in (7).

$$\begin{aligned} x_{k-1} &= c^{-1} y_{k-1} - c^{-1} d w_{k-1} \\ x_{k-2} &= c^{-1} y_{k-2} - c^{-1} d w_{k-2} \end{aligned} \tag{7}$$

Substituting the equation (7) in (5) has (8):

$$y_k = c(a_{1,k-1}(c^{-1} y_{k-1} - c^{-1} d w_{k-1}) + a_{2,k-1}(c^{-1} y_{k-2} - c^{-1} d w_{k-2}) + b v_{k-1}) + d w_k$$

$$y_k = c a_1 c^{-1} y_{k-1} + c_k a_2 c^{-1} y_{k-2} - c a_1 c^{-1} d w_{k-1} - c a_2 c^{-1} d w_{k-2} + c b v_{k-1} + d w_k \tag{8}$$

Reducing terms as shown in (9)

$$y_k = \tilde{a}_1 y_{k-1} + \tilde{a}_2 y_{k-2} + \tilde{d}\tilde{w}_k \tag{9}$$

Where the Gaussian general noise (10)

$$\tilde{d}\tilde{w}_k = -a_1\tilde{d}\tilde{w}_{k-1} - a_2\tilde{d}\tilde{w}_{k-2} + cbv_{k-1} + \tilde{d}\tilde{w}_k,\ \tilde{a}_1 = a_1,\ \tilde{a}_2 = a_2 \tag{10}$$

Considering the equation (9) in (4) the identification error has the form (11):

$$e_k = \tilde{a}_1 y_{k-1} + \tilde{a}_2 y_{k-2} + \tilde{d}\tilde{w}_k - \hat{y}_k \tag{11}$$

Developing (11) in (3) obtains the functional error (12):

$$\begin{aligned} J(e_k^2) &= \tilde{a}_1^2 E\{y_{k-1}^2\} + \tilde{a}_2^2 E y_{k-2}^2 + \tilde{d}_k^2 E\{\tilde{w}_k^2\} + E\{\hat{y}_k^2\} - \\ &2[(\tilde{a}_1 E\{\hat{y}_k y_{k-1}\} + \tilde{a}_2 E\{\hat{y}_k y_{k-2}\} + \tilde{d}_k E\{\hat{y}_k \tilde{w}_k\}) + \\ &((\tilde{a}_2\tilde{d})E\{y_{k-2}\tilde{w}_k\} + (\tilde{a}_1\tilde{a}_2)E\{y_{k-1}y_{k-2}\} + (\tilde{a}_1\tilde{d})E\{y_{k-1}\tilde{w}_k\})] \end{aligned} \tag{12}$$

The stochastic gradient of (12) with respect to $\tilde{a}_1$ achieved in (13).

$$\nabla J(e_k^2)_{\tilde{a}_1} = 2\tilde{a}_1 E\{y_{k-1}^2\} - 2[(E\{\hat{y}_k y_{k-1}\}) + (\tilde{a}_2 E\{y_{k-1}y_{k-2}\} + dE\{y_{k-1}\tilde{w}_k\})] \tag{13}$$

From (13) the estimated value $\tilde{a}_1$ considering $E\{y_{k-1}y_{k-2}\} = 0$, $E\{y_{k-1}\tilde{w}_k\} = 0$ is described in (14)

$$\hat{\tilde{a}}_{1,k} = \frac{E\{\hat{y}_k y_{k-1}\})}{E\{y_{k-1}^2\}} \tag{14}$$

The stochastic gradient of (13) with respect to $\hat{\tilde{a}}_2$, is shown in (15)

$$\hat{\tilde{a}}_{2,k} = \frac{E\{\hat{y}_k y_{k-2}\})}{E\{y_{k-2}^2\}} \tag{15}$$

The adaptive identification finally has the form (16)

$$\hat{y}_k = \hat{\tilde{a}}_{1,k} y_{k-1} + \hat{\tilde{a}}_{2,k} y_{k-2} + \tilde{d}\tilde{w}_k \,.\ \blacksquare \tag{16}$$

3. Conclusion

In the black box concept, the exponential transition function is inaccessible considering the unknown internal parameters and the identification does not operate with these conditions.

The estimation technique based on stochastic gradient solves this difficulty. The system considered was a single input - single output (SISO) model with two delays, described as a second order structure. The estimation parameters were built on the second probability moment form and have adaptive strategies allowing a sufficient convergence rate in the gradient sense. Finally, the identification was integrated with the both estimations into a traditional identification. The theoretical result described the filter strategies integrated as an adaptive filter structure.

References

1. J. J. Medel, R. U. Parrazales, R. Palma., "Estimador Estocástico para un sistema tipo caja negra", Revista Mexicana de Física, Vol. 57, No.3 PP. 204-210 (2011) ISSN 0035-001X.
2. V Valerii, S. Fedorov, L. Leonov, "Parameter Estimation for Models with Unknown Parameters in Variances, Communications in Statistics" Theory and Methods, Vol. 33, Issue 11 pp. 2627 – 2657. (2006) DOI: 10.1081/STA-200037917.
3. F. Kwun Wanga, C. Wen Leea, An M-"Estimator for Estimating the Extended Burr Type III Parameters with Outliers Communications in Statistics" Theory and Methods, Vol. 3, Issue 2, pp. 304 – 322. (2010) DOI 10.1080/03610920903411234.
4. D. Belie, F. Melkebeek, "Seamless integration of a low-speed position estimator for IPMSM in a current-controlled voltage-source inverter", Dept. Electr. Energy, Syst. & Autom., G Univ., Ghent, Belgium Vol. 40, Issue 4, pp.50-55. (2010) DOI 10.1109/SLED.2010.5542804.
5. F. Antreich, J Nossek, A. Seco, Gonzalo; L. Swindlehurst, "Time delay estimation in a spatially structured model using decoupled estimators for temporal and spatial parameters " Smart Antennas, WSA International ITG on Digital Object Identifier Vol. 40, Issue 4,PP. 16-20 (2008) DOI 10.1109/WSA.2008.4475531.
6. C. Breithaupt R. Martin, "Analysis of the Decision-Directed SNR Estimator for Speech Enhancement With Respect to Low-SNR and Transient Conditions", Audio Speech, and Language Processing, IEEE Transactions on Vol. 19, Issue: 2 pp. 277-289. (2011) ISSN. 1558-7916.
7. C. Jian-Zhong, J. Huang, C. Zhi-Ming; "Forecasting of target air route based on weighted least squares method" Control Conference 29th Chinese Vol. 40, Issue 4, pp.3144 – 3147. (2010) ISBN 978-1-4244-6263-6.
8. L., Chengyu Cao, N. Hovakimyan, C. Woolsey, H. Guoqiang, "Estimation of an Affine Motion American Control" Conference St. Louis, MO, USA Vol. 40, Issue 4, pp. 5058-5090. (2009) ISSN 0743-1619.
9. J. Douglas L., Swaroop A, B. Matthew, M. Haun, D. Moussa, D. Sachs, "Adaptive Filtering: LMS Algorithm", http://cnx.org/content/m10481/2.13/.

10. J. J. Medel, M. T. Zagaceta, "Estimador de Parámetros para Sistemas de Orden Superior", Revista Mexicana de Física Vol. 58, No.2, pp. 127-132 (2012) ISSN 0035-001X.
11. H. A. Ruiz, E. M. Toro, R. A. Gallego., "Identificación de errores de difícil detección en estimación de estado en sistemas eléctricos usando algoritmos de optimización combinatorial" Rev. Fac. Ing. Univ. Antioquia N.° 56 pp. 182-192. (2010) ISSN 0120-6230.
12. H. Kirshner., S Maggio., and M. Unser., A Sampling Theory Approach for Continuous ARMA Identification, IEEE Transactions on Signal Processing, vol. 59, No 10, (2011) 4620-4633.
13. I. Morales., M. González., Comparación de las técnicas de análisis de variancia y regresión lineal múltiple: aplicación a un experimento de almacenamiento de mango, Centro Nacional de Ciencia y Tecnología de Alimentos, UCR, Costa Rica, vol. 4, No 7, (2003) 78-83.
14. S. Sezginer., P. Bianchi., Asymptotically Efficient Reduced Complexity Frequency Offset and Channel Estimators for Uplink MIMO-OFDMA Systems Signal Processing, IEEE Transactions on Audio, Speech, and Language Processing, vol. 56 , Issue 3, (2008) 964-979.
15. J Rodríguez, M. Gutiérrez, D. Duarte, J. Mendiola, M. Santillán, "Design and Implementation of an Adjustable Speed Drive for Motion Control Applications" Journal of Applied Research and Technology, Vol. 10 No. 12, (2012) pp. 227-241.
16. F. Casco, R. Medina, M. Guerrero, "A New Variable Step Size NLMS Algorithm and its Performance Evaluation Echo Cancelling Applications" Journal of Applied Research and Technology, Vol. 9 No 03 (2011) pp. 302-313.
17. M. Bazdresch, J. Cortez, O. Longoria, R. Parra, "A Family of Hybrid Space-Time Codes for MIMO Wireless Communications" Journal Applied Research and Technology, Vol. 10 No 02 (2012) pp. 122-142.
18. M. Joham , M. D. Zoltowski, "Interpretation of the Multi-Stage Nested Wiener Filter in the Krylov Subspace Framework", School of Electrical Engineering, Purdue University, West Lafayette , (2000) pp. 1-18.
19. M.S. Koo, H.L. Choi, J.T. Lim, "Output feedback regulation of a chain of integrators with an unbounded time-varying delay in the input" Automatic Control, IEEE Transactions Vol. 57 No 10 (2012) pp.2662-2668.
20. J. J. Medel, M. T. Zagaceta. "Estimación–identificación como filtro digital integrado: descripción e implementación recursiva" Revista Mexicana de Física. Vol. 56, México (2010) pp. 1-8.
21. J. J. Medel, J, García, R. Urbieta. "Identificador con comparación entre dos estimadores" Revista Mexicana de Física . Vol. 57, México (2011) pp. 414-420.
22. H Nyquist, Certain Topics in Telegraph Transmisson Theory, AIEE (Dan Lavry of Lavry Engineering Inc 1928).
23. K. Ogatta, Sistemas de Control en Tiempo Discreto 2 (1995).

Modelo de segmentación para la extracción de características de una señal EEG*

José Luis-Paredes, Hiram Calvo, Jesús Figueroa Nazuno

Centro de Investigación en Computación, Instituto Politécnico Nacional,
Av. Juan de Dios Bátiz e/M.O. Mendizábal s/n, Nva. Ind. Vallejo, 07738
{paredes_a13, hcalvo, jfn}@cic.ipn.mx

Abstract. In this work, a new experimental model of feature extraction for EEG signals is proposed, having as main characteristic the generation of patterns that can be classified with less computational cost and better performance. The proposed algorithm is based on segmentation stripes of the signal applied to a test consisting on recalling the name of one color out of 4, that are presented to a subject of study in screen. During this period the EEG activity of the subject is measured with an EMOTIV device. The proposed segmentation method was applied to 100 samples to classify them according to the corresponding color, and then 9 different machine learning techniques were used. In 8 of them, the proposed model performed better; the simplicity of our model helps to achieve better performance in space and time at the time of classification.

1 Introducción

Desde sus inicios hace poco más de 30 años las Interfaces Cerebro - Computadora o BCI (por sus siglas en inglés) se han enfrentado a diversos retos de investigación y desarrollo, debido principalmente a cuestiones técnicas, análisis de información y entendimiento sobre el funcionamiento del cerebro humano. Hoy en día existen diversos estudios alrededor de este problema, por ejemplo: Imagen por Resonancia Magnética (fMRI) o Magneto encefalografía (MEG) [1, 2]; sin embargo para estos últimos se requiere de equipos altamente sofisticados con altos costos y laboratorios con condiciones adecuadas. La opción más viable para muchos de los investigadores desde los inicios hasta la fecha sigue siendo la Electroencefalografía (EEG) [3, 4]. La EEG requiere de equipos menos costosos que otras interfaces. Dentro de este contexto, los avances tecnológicos han dado paso a una nueva generación de dispositivos portátiles tales como NeuroSky o EMOTIV [5], los cuales han permitido que la investigación y el desarrollo de sistemas BCI sea más accesible. Es importante mencionar que el rendimiento obtenido en comparación con equipos profesionales es menor, debido principalmente a la frecuencia de muestreo y la capacidad de suprimir el ruido de la señal [6, 7, 8].

* Agradecemos el apoyo del gobierno de México y el IPN (CONACYT-SNI, SIP-IPN, COFAA-IPN y BEIFI-IPN)

©L. Berthet, J. C. Chimal, E. A. Santos, E. Castillo
Advances in Computer Science
Research in computing Science 81, 2014 pp. 163-171

Sin importar la finalidad con la que se emplee, un sistema BCI comúnmente tiene los siguientes componentes: adquisición de señal, procesamiento de la señal, extracción de características de la señal e interpretación, como se muestra en la Figura 1.

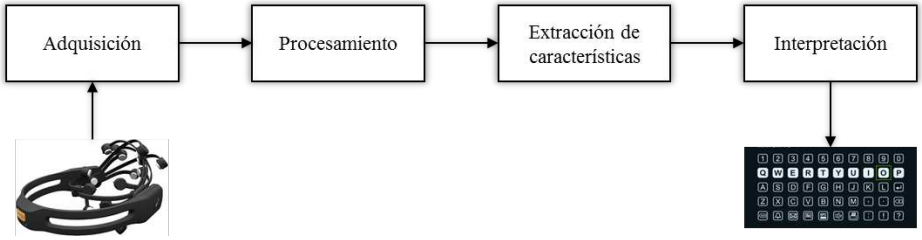


Figura 1.- Componentes de un BCI.

En la actualidad la mayor parte del trabajo de investigación está orientado a los componentes de [9]:

- **Procesamiento**: Se utilizan técnicas tradicionales de análisis de señales principalmente para realizar el filtrado y remoción de artefactos.
- **Extracción de características**: En la mayoría de los casos se usan métodos basados en estadística como son: Análisis de Coeficientes Principales (PCA) o Patrón Espacial Común (CSP), aunque también son utilizadas técnicas tradicionales de procesamiento de señales como: Coeficientes de Predicción Legal (LPC) o *Spectrum*.
- **Interpretación**: En la mayoría de los casos se usan clasificadores de tipo Bayesiano, Logístico, Lineal y Máquinas de Soporte Vectorial (SVM).

Dentro de este contexto tenemos a la utilización de la Potencial Relación – Evento (ERP) [10] como medida de respuesta a estímulos sensoriales y/o cognoscitivos. Este tipo de estímulo está compuesto por factores exógenos que tienen que ver con el entorno físico (sonido, luz, efecto de ruido y corrientes eléctricas que puedan generar estática al momento de obtener la señal), y componentes endógenos, que están relacionados con el estado mental del individuo (estado físico, anímico y de concentración). Con la información generada por el EEG y la utilización de ERP se ha logrado encontrar patrones morfológicos que corresponden a la frecuencia y amplitud de las señales en respuesta a la presentación de estímulos externos y por el efecto de pensamiento o evocación.

El presente trabajo tiene como propósito proponer un nuevo modelo experimental de extracción de características de señales EEG al cual identificaremos como SEG (por segmentación), teniendo como principales características: la generación de datos que puedan ser clasificados más fácilmente y con un menor costo computacional.

2 Metodología experimental

2.1 Adquisición EEG

Se usó un dispositivo EMOTIV el cual consta con 14 electrodos posicionados de acuerdo al sistema 10 – 20 [11] en las siguientes regiones: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8 y AF4 como se muestra en la Figura 2, con una velocidad de muestreo de 128 Hz.

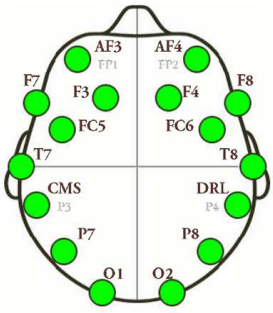


Figura 2.- Ubicación de los electrodos de acuerdo al sistema 10 – 20.

2.2 Procedimiento

El sujeto fue sentado cómodamente frente a un monitor de 14" pulgadas. Antes de comenzar con la adquisición de datos, el sujeto fue sometido a 5 sesiones de práctica con el objetivo de familiarizarlo con la prueba y mantener las condiciones exógenas de manera constante. Posteriormente a las sesiones de prueba el sujeto declaró estar tranquilo y concentrado en la tarea asignada. La prueba consistió en 40 sesiones cortas de 46 segundos con un intervalo de descanso de 60 segundos cada 5 sesiones. Una sesión consta de 12 periodos de inactividad donde se presenta el fondo blanco con una duración de 3 segundos, intercalados con periodos de actividad donde se presenta el fondo con uno de los siguientes colores: "Amarillo, Azul, Rojo, o Verde" con una duración de 1 segundo. Dentro de este periodo, el sujeto hace la evocación del nombre del color en pantalla. Al final de la prueba se tienen 10 segmentos de actividad por cada una de las sesiones. Como resultado se tiene un conjunto de datos compuesto por 100 muestras de cada uno de los colores presentados.

2.3 Análisis EEG

Cada una de las muestras representa al conjunto de datos el cual será analizado de dos formas diferentes: en la primera se tomará el conjunto de datos en bruto para aplicar el modelo propuesto; en el segundo caso, el conjunto de datos será sometido a un filtrado en el dominio de la frecuencia (aplicación de la Transformada Rápida de Fourier (FFT) para generar los coeficientes, mientras que la Transformada Inversa de

Fourier (IFFT) realiza el proceso de filtrado), para posteriormente ser analizados con los siguientes modelos: Análisis de componentes principales (PCA), Coeficientes de predicción lineal (LPC), Densidad espectral de potencia (PSD), Transformada de ondeleta discreta (DWT), Patrón espacial común (CSP). Con cada uno de los análisis realizados se generará un conjunto de datos que representa los patrones que serán sometidos a los siguientes tipos de clasificadores: Bayesianos, Máquinas de Soporte Vectorial (SVM), Regresión Lineal, entre otros. Estos clasificadores fueron elegidos debido a que son los más utilizados de acuerdo al estado del arte [9, 12, 13].

3 Modelo propuesto

El modelo experimental propuesto (SEG) para la extracción de características de una señal EEG (las cuales generalmente se encuentran representada en forma de matrices bidimensionales), consta de 3 etapas las cuales se describen a continuación:

1. Se realiza el proceso de normalización para cada uno de los canales que componen la señal.
2. Se propone un valor para N el cual representara el número de intervalos comprendidos entre 0 y 1. Por ejemplo, si se elige N = 10 entonces tendremos los siguientes intervalos: [0 - .09], [.10 - .19],...., [.9 - 1].
3. Para cada una de las dimensiones de análisis se realiza la sumatoria de cada uno de los valores que se encuentran dentro del intervalo correspondiente. Por ejemplo, si se tienen 5 muestras con los valores (‘.12, .14, .15, .19, .10), los cuales se encuentran dentro del intervalo [.10 - .19], éstos serán sumados obteniendo .70 el cual será asignado como valor final al intervalo anteriormente mencionado.

Este procedimiento se encuentra explicado de manera gráfica en la Figura 3A. En esta podemos ver representada la señal de un solo canal a manera de ejemplo; si hacemos la sumatoria de cada uno de los valores que se encuentran en el intervalo de 0.6 y 0.7 tenemos un valor total de 7.27 el cual se ve reflejado en la Figura 3.B.

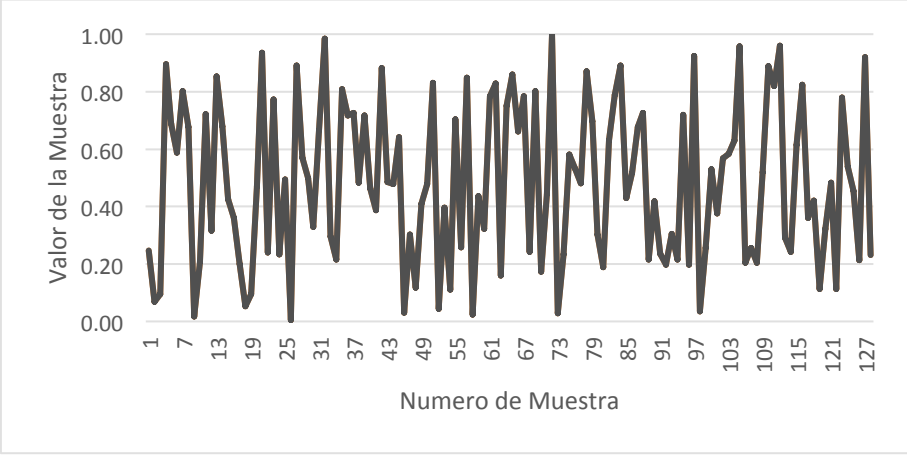


Figura 3A. Señal normalizada para un solo canal.

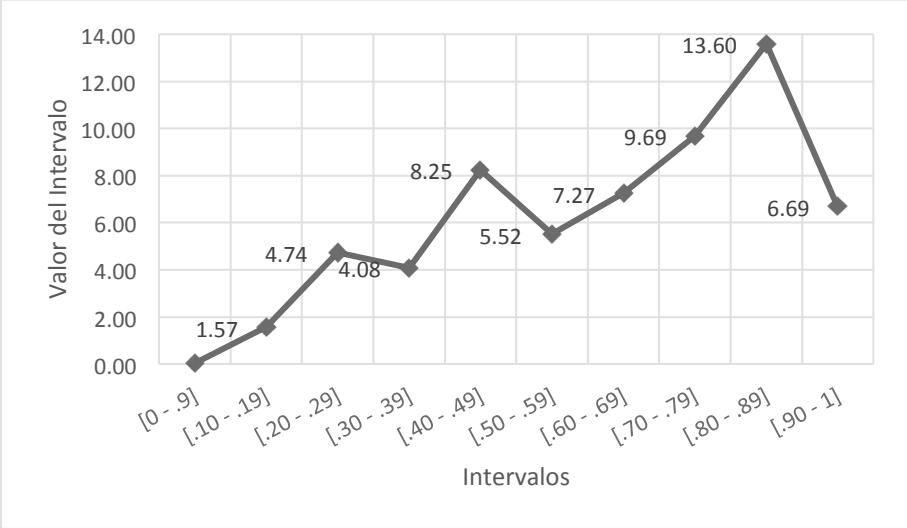


Figura 3B. Resultado del procedimiento de sumatoria para cada uno de los intervalos.

3.1 Pseudocódigo

Entrada: Matriz EEG bidimensional N x M, donde N es el número de canales y M el número de muestras.

Salida: Vector de longitud S x N donde S representa el número de segmentos (o atributos que contendrá el patrón por cada canal.

```
1  Para i ← 1 hasta N hacer
2    Normalizar elementos de i
3  Fin
4  Crear matriz PatronTmp_NxS ← 0
5  Para i ← 1 hasta N hacer
6    Para j ← 1 hasta M hacer
8      Para k ← 1 hasta S hacer
7        PatrónTmp_i,k += Valor_i,j ∈ Segmento
8      Fin Para
9    Fin Para
10 Fin Para
12 Para i ← 1 hasta N hacer
13   Patrón = ConcatHorz(Patrón, PatrónTmp_i)
14 Fin Para
11 Regresar Patrón
```

Algoritmo 1.- Extracción de características de una señal EEG para la generación de un patrón con N atributos.

4 Resultados

El conjunto final de análisis está constituido por 5 conjuntos de datos con 400 instancias, correspondientes a 100 muestras por cada uno de los colores: Amarillo, Azul, Rojo, Verde. Cada uno de los conjuntos de datos fue subdividido de igual manera en 80% para la fase de entrenamiento y 20% para la fase de clasificación. Cada uno fue sometido a clasificación con ayuda de los algoritmos implementados en Weka 3.7.11 sin ningún ajuste en particular para algún conjunto, teniendo como resultados los mostrados en la Tabla 1.

Tabla 1.- Tabla comparativa entre los modelos de extracción de características.

	SEG	CSP	PCA	LPC	PSD
LibSVM	88.75%	70.00%	67.50%	61.25%	55.00%
SimpleLogistic	83.75%	75.00%	68.75%	66.25%	63.75%
SMO	82.50%	73.75%	65.00%	60.00%	50.00%
BayesNet	81.25%	68.75%	61.25%	65.00%	58.75%
LibLINEAR	81.25%	50.00%	36.25%	56.25%	20.00%
RandomSubSpace	80.00%	73.75%	73.75%	66.25%	67.50%
Bagging	77.50%	68.75%	63.75%	56.25%	60.00%
RandomForest	75.00%	75.00%	70.00%	57.50%	56.25%
NNGe	65.00%	71.25%	68.75%	66.25%	58.75%
Promedio	**79.44%**	**69.58%**	**63.89%**	**61.67%**	**54.44%**

Como podemos observar, el modelo propuesto denominado Segmentación (SEG) tiene un mejor desempeño. Cabe mencionar que para este ejercicio en específico se ha elegido una valor de N = 50; sin embargo, al aumentar el número de segmentos se puede obtener un mejor resultado. Bajo el mismo contexto, para la extracción de características dentro del modelo propuesto no se utilizó ningún tipo de filtrado, ya que al momento de hacerlo se tuvo un menor rendimiento. Como ejemplo de ello, para el modelo de clasificación LisbSVM sin filtrado se obtuvo un 88.75% mientras que para el mismo ejercicio, pero con los datos filtrados en el dominio de la frecuencia se obtuvo 63.75%. Las ventajas del modelo propuesto no solo se limitan a obtener un mejor resultado al momento de la clasificación, ya que este resultado lo realiza con un costo computacional inferior con respecto a los demás análisis debido principalmente al número de atributos necesarios para obtener dichos resultados.

Dentro del estado del arte, los resultados obtenidos con diversos métodos de análisis y clasificación son variados, pues existen diversas variables como equipo de adquisición de las señales EEG, condiciones de laboratorio y sujetos de prueba, así como la metodología de la prueba. Todo esto hace poco factible una comparación directa en este tipo de experimentos; sin embargo, en la mayoría de éstos se tiene un porcentaje de clasificación correcta que oscila entre el 60% y 89%. Si tenemos esto como referencia, el obtener puntajes de 65% a 88.75% parece prometedor.

5 Conclusiones

Dentro del estado de arte moderno del estudio de las señales EEG nos podemos dar cuenta de que la mayoría de las técnicas de análisis utilizadas y que presentan mejores resultados son aquellas que están basadas en técnicas estadísticas, debido principalmente a que toman en consideración las pequeñas variaciones entre cada una de las muestras. Esto es fundamental, ya que es allí en donde se encuentra la información que nos interesa en este tipo de señal. Teniendo esto en consideración y con base en los resultados podemos apreciar que los modelos SEG, CSP y PCA (los cuales son meramente estadísticos) pueden realizar una mejor extracción de características para una correcta clasificación en comparación con modelos como LPC y PSD, los cuales están basados en análisis tradicionales de señales.

Dentro del análisis de resultados se observó que el modelo propuesto tiene un mejor desempeño en la mayoría de los casos sin necesidad de hacer un filtrado en el dominio de la frecuencia a comparación de los demás, donde se obtuvo una mejor clasificación, en especial LPC y PSD. En comparación con CSP y PCA, además de obtener una mejor clasificación, el modelo propuesto tiene también como característica un menor costo computacional en tiempo de procesamiento y espacio del patrón generado.

Uno de los factores importantes a considerar es que se puede realizar una variación de la cantidad de características que pueden ser generadas; esto le permite ser un modelo dinámico que puede adaptarse a las frecuencias y amplitudes del tipo de señales a analizar. Si bien este modelo está pensado para tratar problemas multidimensionales, sin embargo también puede ser utilizado para la extracción de características en señales unidimensionales.

Referencias

1. Yoichi Miyawaki, Hajime Uchida, Okito Yamashita, Masaaki Sato, Yusuke Morito, Hiroki C. Tanabe (2008). Visual Image Reconstruction from Human Brain Activity using a Combination of Multiscale Local Image Decoders. Neuron 60, 915 –929.
2. Alexander M. Chan, Eric Halgren, Ksenija Marinkovic, Sydney S. Cash (2011). Decoding word and category-specific spatiotemporal representations from MEG and EEG. NeuroImage 54 (2011) 3028–3039.
3. Saeid Sanei y J.A. Chambers (2007). EEG Signal Processing. John Wiley & Sons Ltd.
4. Shanbao Tong, Nitish V. Thakor (2009). Quantitative EEG Analysis Methods and Clinical Applications. Engineering in Medicine & Biology. Artech House
5. Emotiv EPOC. http://www.emotiv.com
6. Nicholas A. Badcock, Petroula Mousikou, Yatin Mahajan, Peter de Lissa, Johnson Thie y Genevieve McArthur (2013). Validation of the Emotiv EPOC

EEG gaming systemfor measuring research quality auditory ERPs. PeerJ, DOI 10.7717/peerj.38
7. Matthieu Duvinage, Thierry Castermans, Mathieu Petieau, Thomas Hoellinger, Guy Cheron and Thierry Dutoit (2013). Performance of the Emotiv Epoc headset for P300-based applications. BioMedical Engineering OnLine 2013.
8. Athanasios Vourvopoulos, Fotis Liarokapis (2014). Evaluation of commercial brain – computer interfaces in real and virtual world environment: A pilot study. Computers and Electrical Engineering 40, 714–729.
9. Xing-Yu Wang, Jing Jin, Yu Zhang, Bei Wang (2013). Brain Control: Human-computer Integration Control Based on Brain-computer Interface Approach. Acta Automatica Sinica, 2013, Vol. 39, No. 3.
10. Marta Kutas y Kara D. Federmeier (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). Annu. Rev. Psychol. 2011.62:621-647.
11. Jasper, H. H., The ten-twenty electrode system of the International Federation. Electroen. Clin. Neuro. 10, 371–375, 1958.
12. Nikolay V. Manyakov, Nikolay Chumerin, Adrien Combaz, y Marc M. Van Hulle (2011). Comparison of Classification Methods for P300 Brain-Computer Interface on Disabled Subjects. Computational Intelligence and Neuroscience Volume 2011, Article ID 519868.
13. Jason Farquhar y N. Jeremy Hill (2012). Interactions between Pre-Processing and Classification Methods for Event-Related-Potential Classification: Best-Practice Guidelines for Brain-Computer Interfacing. Neuroinformatics April 2013, Volume 11, Issue 2, pp 175-192.

Reviewing Committee

Alejo Macías Miguel Ángel
Altamirano Álvaro
Argüelles Cruz Amadeo José
Barrón Fernández Ricardo
Bonilla Licea Daniel
Botello Castillo Alejandro
Calvo Castro Francisco Hiram
Carreto Arellano Chadwick
Castillo Montiel Erandi
Chimal Eguía Juan Carlos
Cleofas Sánchez Laura
Cruz Cortés Nareli
Delgado Hernández Julio Carlos
Felipe Riverón Edgardo Manuel
Figueroa Nazuno Jesús Guillermo
Godoy Calderón Salvador
Guevara Martínez Elizabeth
Gutierrez Aldana Alfonso
Gutierrez García Juan Jesús
Guzmán Lugo José Giovanni
Horacio Dominguez Christian
Huerta Trujillo Iliac
Juárez Gambino Joel Omar
Landassuri Moreno Victor Manuel
Lazcano Salas Saul
López Ramírez Blanca Cecilia
Martínez Ibarra Mario Iván
Martínez Luna Gilberto Lorenzo
Márquez de Silva Sergio Antonio
Márquez Molina Miguel
Menchaca Méndez Rolando
Mendoza Mendoza Julio Alberto
Miranda Jiménez Sabino
Morales Escobar Saturnino Job
Moreno Armendáriz Marco Antonio
Navarrete Manzanilla Niels Henrik
Orantes Jiménez Sandra Dinora
Orozco Aguirre Héctor Rafael
Pineda Briseño Anabel
Quintana López Maricela
Rubio Espino Elsa
Ruiz Ibañez Victor Antonio
Salinas Rosales Moisés
Sidorov Grigori Olegovich
Sossa Azuela Juan Humberto
Suárez Guerra Sergio
Tamariz Flores Edna Iliana
Téllez Castillo Germán
Torrealba Meléndez Richard
Valle Chavez Abel
Yáñez Márquez Cornelio

Impreso en los Talleres Gráficos
de la Dirección de Publicaciones
del Instituto Politécnico Nacional
Tresguerras 27, Centro Histórico, México, D.F.
noviembre de 2014
Printing 500 / Edición 500 ejemplares

ISSN: 1870-4069
http://rcs.cic.ipn.mx

RCS
Research in Computing Science